(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 September 2003 (04.09.2003)

PCT

(10) International Publication Number
WO 03/072041 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number: PCT/US03/05458

(22) International Filing Date: 23 February 2003 (23.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/360,274 27 February 2002 (27.02.2002) US

(71) Applicant *(for all designated States except US)*: **MERCK & CO., INC.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ESPESETH, Amy, S.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **FERRER, Marc** [ES/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **FLORES, Osvaldo, A.** [CL/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **HAZUDA, Daria, J.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **INGLESE, James** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **MILLER, Michael, D.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **REGISTER, Bruce** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **SHI, Xiao-Ping** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **SIMON, Adam, J.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **ZUCK, Paul, D.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

(74) Common Representative: **MERCK & CO., INC.**; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

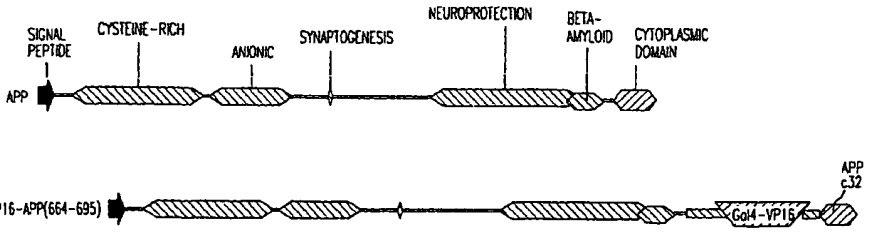
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ASSAYS TO MONITOR AMYLOID PRECURSOR PROTEIN PROCESSING



(57) Abstract: The present invention provides DNA constructs, genetically engineered host cells, and methods for identifying inhibitors of amyloid precursor protein (APP) processing. The methods provide for the convenient identification, in a single assay, of inhibitors of β-secretase and γ-secretase as well as other forms of APP processing. The methods rely on fusion proteins of APP and transcription factors in which APP processing releases the transcription factors, allowing the transcription factors to activate transcription of a reporter gene. Inhibitors are identified as substances that block or diminish transcription factor release from the fusion protein, thereby causing a diminution of reporter gene readout.

WO 03/072041 A2

TITLE OF THE INVENTION

ASSAYS TO MONITOR AMYLOID PRECURSOR PROTEIN PROCESSING

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 60/360,274, filed February 27, 2002, the contents of which are incorporated herein by reference in their entirety.

STATEMENT REGARDING FEDERALLY-SPONSORED R&D

Not applicable.

REFERENCE TO MICROFICHE APPENDIX

Not applicable.

FIELD OF THE INVENTION

The present invention is directed to the field of Alzheimer's disease. In particular, the present invention provides novel methods of identifying substances that are specific inhibitors of various steps in the processing of amyloid precursor protein.

BACKGROUND OF THE INVENTION

Alzheimer's disease is a common, chronic neurodegenerative disease, characterized by a progressive loss of memory and sometimes severe behavioral abnormalities, as well as an impairment of other cognitive functions that often leads to dementia and death. It ranks as the fourth leading cause of death in industrialized societies after heart disease, cancer, and stroke. The incidence of Alzheimer's disease is high, with an estimated 2.5 to 4 million patients affected in the United States and perhaps 17 to 25 million worldwide. Moreover, the number of sufferers is expected to grow as the population ages.

A characteristic feature of Alzheimer's disease is the presence of large numbers of insoluble deposits, known as amyloid plaques, in the brains of those affected. Autopsies have shown that amyloid plaques are found in the brains of virtually all Alzheimer's patients and that the degree of amyloid plaque deposition correlates with the degree of dementia (Cummings & Cotman, 1995, Lancet

326:1524-1587). While some opinion holds that amyloid plaques are a late stage by-product of the disease process, the consensus view is that amyloid plaques are more likely to be intimately, and perhaps causally, involved in Alzheimer's disease.

A variety of experimental evidence supports this view. For example, Aβ, a primary component of amyloid plaques, is toxic to neurons in culture and transgenic mice that overproduce Aβ in their brains show significant deposition of Aβ into amyloid plaques as well as significant neuronal toxicity (Yankner, 1990, Science 250:279-282; Mattson et al., 1992, J. Neurosci. 12:379-389; Games et al., 1995, Nature 373:523-527; LaFerla et al., 1995, Nature Genetics 9:21-29). Mutations in the APP gene, leading to increased Aβ production, have been linked to heritable forms of Alzheimer's disease (Goate et al., 1991, Nature 349:704-706; Chartier-Harlan et al., 1991, Nature 353:844-846; Murrel et al., 1991,Science 254:97-99; Mullan et al., 1992, Nature Genetics 1:345-347). Presenilin-1 (PS1) and presenilin-2 (PS2) related familial early-onset Alzheimer's disease (FAD) shows disproportionately increased production of Aβ1–42, the 42 amino acid isoform of Aβ, as opposed to Aβ1–40, the 40 amino acid isoform (Scheuner et al, 1996, Nature Medicine 2:864–870). The longer isoform of Aβ is more prone to aggregation than the shorter isoform (Jarrett et al, 1993, Biochemistry 32:4693-4697). Injection of the insoluble, fibrillar form of Aβ into monkey brains results in the development of pathology (neuronal destruction, tau phosphorylation, microglial proliferation) that closely mimics Alzheimer's disease in humans (Geula et al., 1998, Nature Medicine 4:827-831). See Selkoe, 1994, J. Neuropathol. Exp. Neurol. 53:438-447 for a review of the evidence that amyloid plaques have a central role in Alzheimer's disease.

Aβ, a 39-43 amino acid peptide derived by proteolytic cleavage of the amyloid precursor protein (APP), is the major component of amyloid plaques (Glenner & Wong, 1984, Biochem. Biophys. Res. Comm. 120:885-890). APP is actually a family of polypeptides produced by alternative splicing from a single gene. Major forms of APP are known as APP_{695}, APP_{751}, and APP_{770}, with the subscripts referring to the number of amino acids in each splice variant (Ponte et al., 1988, Nature 331:525-527; Tanzi et al., 1988, Nature 331:528-530; Kitaguchi et al., 1988, Nature 331:530-532). APP is membrane bound and undergoes proteolytic cleavage by at least two pathways. In one pathway, cleavage by an enzyme known as α-secretase occurs while APP is still in the trans-Golgi secretory compartment (Kuentzel et al., 1993, Biochem. J. 295:367-378). This cleavage by α-secretase

occurs within the Aβ portion of APP, thus precluding the formation of Aβ. In another proteolytic pathway, cleavage of the Met596-Asp597 bond (numbered according to the 695 amino acid protein) by an enzyme known as β-secretase occurs. This cleavage by β-secretase generates the N-terminus of Aβ. The C-terminus is formed by cleavage by a second enzyme known as γ-secretase. The C-terminus is actually a heterogeneous collection of cleavage sites rather than a single site since γ-secretase activity occurs over a short stretch of APP amino acids rather than at a single peptide bond. Peptides of 40 or 42 amino acids in length (Aβ1–40 and Aβ1–42, respectively) predominate among the C-termini generated by γ-secretase. Aβ1–42 is more prone to aggregation than Aβ1–40, is the major component of amyloid plaque (Jarrett et al., 1993, Biochemistry 32:4693-4697; Kuo et al., 1996, J. Biol. Chem. 271:4077-4081), and its production is closely associated with the development of Alzheimer's disease (Sinha & Lieberburg, 1999, Proc. Natl. Acad. Sci. USA 96:11049-11053). The bond cleaved by γ-secretase appears to be situated within the transmembrane domain of APP. It is unclear as to whether the C-termini of Aβ1–40 and Aβ1–42 are generated by a single γ-secretase protease with sloppy specificity or by two distinct proteases. For a review that discusses APP and its processing, see Selkoe, 1998, Trends Cell. Biol. 8:447-453.

Much interest has focused on the possibility of inhibiting the development of amyloid plaques as a means of preventing or ameliorating the symptoms of Alzheimer's disease. To that end, a promising strategy is to inhibit the activity of β- and γ-secretase, the two enzymes that together are responsible for producing Aβ. This strategy is attractive because, if the formation of amyloid plaques as a result of the deposition of Aβ is a cause of Alzheimer's disease, inhibiting the activity of one or both of the two secretases would intervene in the disease process at an early stage, before late-stage events such as inflammation or apoptosis occur. Such early stage intervention is expected to be particularly beneficial (see, *e.g.*, Citron, 2000, Molecular Medicine Today 6:392-397).

To that end, various assays have been developed that are directed to the identification of compounds that may interfere with the production of Aβ or its deposition into amyloid plaques. U.S. Patent No. 5,441,870 is directed to methods of monitoring the processing of APP by detecting the production of amino terminal fragments of APP. U.S. Patent No. 5,605,811 is directed to methods of identifying inhibitors of the production of amino terminal fragments of APP. U.S. Patent No.

5,593,846 is directed to methods of detecting soluble Aβ by the use of binding substances such as antibodies. Esler et al., 1997, Nature Biotechnology 15:258-263 described an assay that monitored the deposition of Aβ from solution onto a synthetic analogue of an amyloid plaque. The assay was suitable for identifying compounds that could inhibit the deposition of Aβ. However, this assay is not suitable for identifying substances, such as inhibitors of β- or γ-secretase, that would prevent the formation of Aβ.

Various groups have cloned and sequenced cDNA encoding a protein that is believed to be β-secretase (Vassar et al., 1999, Science 286:735-741; Hussain et al., 1999, Mol. Cell. Neurosci. 14:419-427; Yan et al., 1999, Nature 402:533-537; Sinha et al., 1999, Nature 402:537-540; Lin et al., 2000, Proc. Natl. Acad. Sci. USA 97:1456-1460) but the identity of γ-secretase has been more elusive. A pair of proteins known as presenilin-1 and presenilin-2 are viewed as possible candidates (Selkoe & Wolfe, 2000, Proc. Natl. Acad. Sci. USA 97:5690-5692).

Presenilin-1 (PS1) and presenilin-2 (PS2) are polytopic membrane proteins that are involved in γ-secretase-mediated processing of APP. The most common cause of familial early-onset Alzheimer's disease is the autosomal dominant inheritance of assorted mutations in the PS1 gene (Sherrington et al., 1995, Nature 375:754–760). These PS1 mutations lead to increased production of Aβ1-42 (Scheuner et al., 1996, Nature Medicine 2:864–870; Duff et al., 1996, Nature 383:710–713; Borchelt et al., 1996, Neuron 17:1005–1013). Similarly, certain mutations in PS2 cause familial early-onset Alzheimer's disease and increased generation of Aβ42 (Levy-Lahad et al., 1995, Science 269:970–973). Cultured isolated neurons from PS1-deficient mice exhibit reduced γ-secretase-mediated cleavage of APP (De Strooper et al., 1998, Nature 391:387–390). It was suggested that PS1 might influence trafficking of APP and/or γ-secretase or it might play a more direct role in proteolytic cleavage of APP. Directed mutagenesis of two conserved transmembrane-situated aspartates in PS1 was shown to inactivate γ-secretase activity in cellular assays, suggesting that PS1 is either a required diaspartyl cofactor for γ-secretase or is itself γ-secretase (Wolfe et al., 1999, Nature 398:513–517). Moreover, Li et al., 2000, Nature 405:689-694 made photoactivatable derivatives of a highly specific and potent aspartyl protease transition state analog inhibitor and found that the inhibitor selectively labeled presenilin fragments.

Co-immunoprecipitation experiments have shown that PS1 and PS2 interact directly with the immature forms of APP in the endoplasmic reticulum where the disease-associated amyloid Aβ1-42 peptide is probably generated (Xia et al., 1997 Proc. Natl. Acad. Sci. USA 94:8208-8213; Weidemann et al., 1997, Nat. Med. 3:328-332). Knock-out of PS1 activity greatly diminishes γ-secretase cleavage of APP (De Strooper et al., 1998, Nature 391:387-390). PS1 knock-outs do not exhibit total lack of γ-secretase activity but knock-out of both PS1 and PS2 activity does result in a total loss of γ-secretase activity (Herreman et al., 2000, Nat. Cell. Biol. 2:461-462; Zhang et al., 2000, Nat. Cell Biol. 2:463-465), suggesting that PS2 has a similar function to PS1 in the processing of APP.

Karlström et al., (Journal of Biological Chemistry papers in press, published on December 13, 2001 as Manuscript C100649200) describes an assay designed specifically to identify inhibitors of γ-secretase cleavage of APP. The authors inserted the GAL4 DNA binding domain fused to the VP16 transactivation domain into C99, a portion of APP containing the 99 carboxy-terminal amino acids. This fragment of APP contains the γ-secretase cleavage site but lacks the β-secretase cleavage site. Transaction of a UAS reporter plasmid by GAL4-VP16 confirmed cleavage of the Gal4-VP16/C99 substrate by γ-secretase only. Thus, the assay is capable of detecting γ-secretase inhibitors but not inhibitors of β-secretase or other modulators of APP processing requiring the N-terminal domain of APP.

Cao & Südhoff, 2001, Science 293:115-120 described work in which the GAL4 and LexA DNA binding domains were inserted into APP to demonstrate the potential of the cleaved C-terminus of APP for transcriptional co-activation. In this article, a transcriptional factor was not fused to APP and no attempt was made to develop an assay for the identification of APP processing inhibitors.

Sisodia, 1992, Proc. Natl. Acad. Sci. USA 89:6975-6979 described various changes in the amino acid sequence of APP in the region of the α-secretase cleavage site and the effect of those changes on cleavage by α-secretase. A change of K to V at position 612 of the 695 amino acid version of APP led to reduced cleavage by α-secretase.

U.S. Patent No. 6,333,167 B1 discloses an assay involving DNA constructs encoding portions of membrane proteins containing sites that are susceptible to cleavage by proteases that are fused to transcriptional repressors. Such constructs are introduced into cells that contain a reporter gene under the control of a

promoter that is sensitive to the repressor. In the absence of an inhibitor of the protease, the fusion protein is cleaved by the protease, releasing a membrane protein/repressor fusion protein that translocates to the nucleus and represses transcription from the reporter gene. In the presence of an inhibitor of the protease, the membrane protein/repressor fusion protein is not released and thus cannot repress transcription from the reporter. An increase in reporter expression can therefore be used as a readout for the presence of an inhibitor.

SUMMARY OF THE INVENTION

The present invention is directed to methods of identifying inhibitors of the processing of amyloid precursor protein (APP) that are capable of identifying inhibitors of a number of steps of such processing. Unlike prior methods, the methods of the present invention can be used to screen for inhibitors of β-secretase cleavage, γ-secretase cleavage, APP extracellular signaling, or APP cytoplasmic signaling in a single assay.

The methods employ a recombinant eukaryotic cell that is capable of processing APP. The cell has been engineered to express a fusion protein that contains amino acid sequences encompassing both the β-secretase cleavage site of APP and the γ-secretase cleavage site. The fusion protein also contains a transcription factor fused in frame to the APP sequences.

When the recombinant cell is further engineered to contain a reporter gene, in which transcription of the reporter gene is driven by a regulatory DNA sequence that is inactive in the absence of the transcription factor but active in the transcription factor's presence, a system useful for screening for APP processing inhibitors is provided. Since the recombinant cell has been selected so as to be capable of processing APP, the fusion protein will be processed, releasing the transcription factor and activating transcription of the reporter gene. The reporter gene has been preselected so that activation of the reporter gene leads to a detectable phenotype.

The system is utilized by exposing the recombinant cell to substances that are to be tested for the ability to inhibit APP processing. Those substances that are actually inhibitors of APP processing will cause diminished processing of the fusion protein, leading to smaller amounts of the transcription factor being released.

This leads to less transcription of the reporter gene. This results in a decrease in the phenotypic effect of the reporter gene that can be observed.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A-G shows a schematic diagram of several APP/transcription factor fusion constructs.

Figure 2A-B shows the DNA sequence (SEQ ID NO:1) of the fusion protein APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695).

Figure 3 shows the amino acid sequence (SEQ ID NO:2) of the fusion protein APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = K612V mutation; 4 = region of γ-secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 4A-B shows the DNA sequence (SEQ ID NO:3) of the fusion protein APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695).

Figure 5 shows the amino acid sequence (SEQ ID NO:4) of the fusion protein APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = K612V mutation; 4 = region of γ-secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 6A-C shows the DNA sequence (SEQ ID NO:5) of the fusion protein APP(1-651)SW, K612V, GAL4-VP16(delMet) APP (664-695).

Figure 7 shows the amino acid sequence (SEQ ID NO:6) of the fusion protein APP(1-651)SW, K612V, GAL4-VP16(delMet) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = K612V mutation; 4 = region of γ-secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 8A-C shows the DNA sequence (SEQ ID NO:7) of the fusion protein APP(1-651)wt, K612V, GAL4-VP16(del Met) APP (664-695).

Figure 9 shows the amino acid sequence (SEQ ID NO:8) of the fusion protein APP(1-651)wt, K612V, GAL4-VP16(del Met) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = K612V mutation; 4 = region of γ-

secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 10A-B shows the DNA sequence (SEQ ID NO:9) of the fusion protein APP(1-651)SW, TATexonI(M1L) APP (664-695).

Figure 11 shows the amino acid sequence (SEQ ID NO:10) of the fusion protein APP(1-651)SW, TATexonI(M1L) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = wild-type K at position 612; 4 = region of γ-secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 12A-B shows the DNA sequence (SEQ ID NO:11) of the fusion protein APP(1-651)wt, TATexonI(M1L) APP (664-695).

Figure 13 shows the amino acid sequence (SEQ ID NO:12) of the fusion protein APP(1-651)wt, TATexonI(M1L) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = wild-type K at position 612; 4 = region of γ-secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 14A-C shows the DNA sequence (SEQ ID NO:13) of the fusion protein APP(1-651)SW, GAL4-VP16(delMet) APP (664-695).

Figure 15 shows the amino acid sequence (SEQ ID NO:14) of the fusion protein APP(1-651)SW, GAL4-VP16(delMet) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = wild-type K at position 612; 4 = region of γ-secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 16A-C shows the DNA sequence (SEQ ID NO:15) of the fusion protein APP(1-651)wt, GAL4-VP16(delMet) APP (664-695).

Figure 17 shows the amino acid sequence (SEQ ID NO:16) of the fusion protein APP(1-651)wt, GAL4-VP16(delMet) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = wild-type K at position 612; 4 = region of γ-secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 18A-B shows the cDNA sequence (SEQ ID NO:17) and Figure 18C shows the amino acid sequence (SEQ ID NO:18) of the 695 amino acid splice variant of wild-type Alzheimer's precursor protein (APP). See GenBank accession no. Y00264 and Kang et al., 1987, Nature 325:733-736.

Figure 19 shows data from an embodiment in which the assay of the present invention was used to identify both a β-secretase inhibitor and a γ-secretase inhibitor. See Example 3 for details.

Figure 20 shows a schematic diagram of pCR2.1 Gal4-VP16.

Figure 21A shows a schematic diagram of pRBR121. Figure 21B shows a schematic diagram of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695).

Figure 22A shows a schematic diagram of pRBR186. Figure 22B shows a schematic diagram of the viral plasmid pNL4-3. Figure 22C shows a schematic diagram of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) with additional details as compared to Figure 21B, which shows the same plasmid.

Figure 23 shows a schematic diagram of pRSV Kan/Neo res.

Figure 24 shows a schematic diagram of pUCd5TAT.

Figure 25A shows a schematic diagram of pMM321. Figure 25B-D shows the nucleotide sequence of pMM321. The upper strand is SEQ ID NO:19. The lower strand (SEQ ID NO:20) is the reverse complement of SEQ ID NO:19.

Figure 26A shows a schematic diagram of the expression vector pcDNA3.1 zeo (+)APP(1-651)SW, K612V-(M1L)TATexonI. This expression vector directs the expression of a fusion protein containing the first 651 amino acids of APP with the Swedish version of the β-secretase cleavage site and the K612V mutation fused to the first exon of HIV1 TAT. The methionine at position 1 of TAT has been changed to leucine. Figure 26B-G shows the nucleotide sequence of pcDNA3.1 zeo (+)APP(1-651)SW, K612V-(M1L)TATexonI. The upper strand is SEQ ID NO:21. The lower strand (SEQ ID NO:22) is the reverse complement of SEQ ID NO:21.

Figure 27A-B shows a schematic diagram depicting general features of the present invention. Figure 27A: The vertical bar represents a fusion protein with APP sequences represented as unfilled or lightly shaded portions of the bar. The lightly shaded portion represents Aβ. "BACE" indicates the β-secretase cleavage site. The dark shaded portion represents the transcription factor fused between APP

sequences. The horizontal bar represents a membrane in which the uncleaved fusion protein is embedded, e.g., the endoplasmic reticulum. Figure 27B: The transcription factor (plus small amounts of APP), having been released from the fusion protein and thus the membrane by APP processing, is shown in the nucleus binding to and activating the regulatory DNA sequence ("Transcription Factor Response Element") that controls the expression of the reporter gene.

Figure 28A-B shows the DNA sequence (SEQ ID NO:23) of the fusion protein APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695).

Figure 29 shows the amino acid sequence of a fusion protein (APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695)) (SEQ ID NO:24) containing the sequence NFEV at the β-secretase cleavage site (underlined at 2). The other portions of the fusion protein are indicated as follows: 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = K612V mutation; 4 = region of γ-secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 30A-C shows the DNA sequence (SEQ ID NO:25) of the fusion protein APP(1-651)NFEV, K612V, GAL4-VP16(delMet) APP (664-695).

Figure 31 shows the amino acid sequence of a fusion protein (APP(1-651)NFEV, K612V, GAL4-VP16(delMet) APP (664-695)) (SEQ ID NO:26) containing the sequence NFEV at the β-secretase cleavage site (underlined at 2). The other portions of the fusion protein are indicated as follows: 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = K612V mutation; 4 = region of γ-secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 32A shows a schematic diagram of pcDNA3.1 zeo (+), a eukaryotic expression vector that is suitable for use in the present invention. Figure 32B-F shows the nucleotide sequence of pcDNA3.1 zeo (+). The upper strand is SEQ ID NO:27. The lower strand (SEQ ID NO:28) is the reverse complement of SEQ ID NO:27.

Figure 33 shows data from an embodiment of the present invention utilizing a β-galactosidase reporter gene in which the assay of the present invention was used to identify both a β-secretase inhibitor and a γ-secretase inhibitor. See Example 8 for details.

Figure 34 shows data from an embodiment of the present invention in which a fusion protein having a wild-type β-secretase cleavage site and a fusion

protein having a Swedish β-secretase cleavage site are compared. See Example 9 for details.

Figure 35A-B shows the DNA sequence (SEQ ID NO:29) of the fusion protein APP(1-651)NFEV, TATexonI(M1L) APP (664-695).

Figure 36 shows the amino acid sequence of a fusion protein (APP(1-651)NFEV, TATexonI(M1L) APP (664-695)) (SEQ ID NO:30) containing the sequence NFEV at the β-secretase cleavage site (underlined at 2) and a wild-type K at position 612 (underlined at 3). The other portions of the fusion protein are indicated as follows: 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = wild-type K; 4 = region of γ-secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 37A-C shows the DNA sequence (SEQ ID NO:31) of the fusion protein APP(1-651)NFEV, GAL4-VP16(delMet) APP (664-695).

Figure 38 shows the amino acid sequence of a fusion protein (APP(1-651)NFEV, GAL4-VP16(delMet) APP (664-695)) (SEQ ID NO:32) containing the sequence NFEV at the β-secretase cleavage site (underlined at 2) and a wild-type K at position 612 (underlined at 3). The other portions of the fusion protein are indicated as follows: 1 = amino acids 1-651 of APP; 2 = region of β-secretase cleavage; 3 = wild-type K; 4 = region of γ-secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

DETAILED DESCRIPTION OF THE INVENTION

For the purposes of this invention:

A "fusion protein" is a protein that contains at least two polypeptide regions and, optionally, a linking peptide to operatively link the two polypeptides into one continuous polypeptide. The at least two polypeptide regions in a fusion protein are derived from different sources, and therefore a fusion protein comprises two polypeptide regions not normally joined together in nature.

A "linking sequence (or linker peptide)" contains one or more amino acid residues joined in peptide bonds. A linking sequence serves to join two polypeptide regions of differing origins in a fusion protein via a peptide bond between the linking sequence and each of the polypeptide regions.

Typically, a fusion protein is synthesized as a continuous polypeptide in a recombinant host cell which contains an expression vector comprising a

nucleotide sequence encoding the fusion protein where the different regions of the fusion protein are fused in frame on either side of a linker peptide's coding sequence. The chimeric coding sequence (encoding the fusion protein) is operatively linked to expression control sequences (generally provided by the expression vector) that are functional in the recombinant host cell.

"Reporter gene," as used in the present invention, does not mean a DNA sequence present on the chromosome of a cell, generally possessing introns, as is often meant by the word "gene" in the art. Rather "reporter gene" means any DNA sequence encoding a protein or polypeptide that can give rise to a signal that can be detected or measured. "Reporter gene" does not mean a portion of the amino acid sequence of APP. "Reporter gene" will usually mean a DNA sequence, generally a cDNA sequence (although in some cases a reporter gene may have introns) that encodes a protein or polypeptide that is commonly used in the art to provide a measurable phenotype that can be distinguished over background signals.

A "nuclear localization signal (NLS)" is a region of a polypeptide which targets the polypeptide to the nucleus of the cell. One such NLS is that from the SV40 large T antigen. See, e.g., U.S. Patent No. 5,589,392; Kalderon et al., 1984, Cell 39:499-509. The minimum region of the SV40 large T antigen with NLS activity is Pro-Lys-Lys-Lys-Arg-Lys-Val (SEQ ID NO:22). See also U.S. Patent No. 5,776,689.

"Substances" that are screened in the present invention can be any substances that are generally screened in the pharmaceutical industry during the drug development process. For example, substances may be low molecular weight organic compounds (*e.g.*, having a molecular weight of less than about 2,000 daltons and preferably less than about 1,000 daltons), RNA, DNA, antibodies, peptides, or proteins. Substances are often tested in the methods of the present invention as large collections of substances, *e.g.* libraries of low molecular weight organic compounds, peptides, or natural products.

The conditions under which substances are employed in the methods described herein are conditions that are typically used in the art for the study of protein-ligand interactions or enzyme inhibition studies: *e.g.*, salt conditions such as those represented by such commonly used buffers as PBS or in tissue culture media; a temperature of about 4°C to about 55°C; incubation times of from several seconds to several hours or even up to 24 or 48 hours. Screening for the identification of

enzyme-specific inhibitors is a well-known procedure in the pharmaceutical arts and the numerous conditions under which such screening has been done are available in the literature to guide the practitioner of he present invention.

A "conservative amino acid substitution" refers to the replacement of one amino acid residue by another, chemically similar, amino acid residue. Examples of such conservative substitutions are: substitution of one hydrophobic residue (isoleucine, leucine, valine, or methionine) for another; substitution of one polar residue for another polar residue of the same charge (*e.g.*, arginine for lysine; glutamic acid for aspartic acid); substitution of one aromatic amino acid (tryptophan, tyrosine, or phenylalanine) for another.

"Transfection" refers to any of the methods known in the art for introducing DNA into a cell, *e.g.*, calcium phosphate or calcium chloride mediated transfection, electroporation, infection with a retroviral vector.

The present invention relates to the discovery of an assay system that permits the simultaneous screening for inhibitors of several types of amyloid precursor protein (APP) processing or signaling (e.g., β-secretase cleavage, γ-secretase cleavage, APP extracellular signaling, APP cytoplasmic signaling). In a preferred embodiment, this screening is accomplished without the concomitant identification of inhibitors of α-secretase. The assay system is carried out in a single type of cell, using a single type of assay readout. Inhibitors discovered by means of the present invention are expected to be useful in the treatment of Alzheimer's disease since these inhibitors are likely to be capable of interfering with the production of Aβ.

Previous assays for identifying inhibitors of APP processing have focussed specifically on inhibition of either β-secretase or γ-secretase activity, or on inhibition of some other single aspect of Aβ production. In contrast, the assays described herein are directed to inhibition of APP processing in general. Substances identified through these assays may target β-secretase, γ-secretase, modulators of β-secretase or γ-secretase activity, or even an as-yet-undiscovered ligand interaction with APP. In certain embodiments, these assays will also be free of the potentially misleading or obscuring effects of α-secretase activity. In addition, unlike other assays currently in use, these assays are homogeneous assays; i.e., they require no cumbersome or time-consuming steps such as column chromatography separations, immunoprecipitations, washing steps, etc. Therefore, the assays are very well adapted to a high throughput screening format.

In the present invention, novel recombinant DNA molecules are constructed in which nucleotide sequences encoding at least a portion of the luminal (i.e., N-terminal to the transmembrane region) and transmembrane regions of APP are fused to nucleotide sequences encoding a transcription factor. In a preferred embodiment, the APP contains an α-secretase cleavage site that has been altered to reduce or eliminate α-secretase cleavage. This allows the assays of the present invention to avoid identifying inhibitors of α-secretase and permits the more efficient detection of β-secretase inhibitors since α-secretase and β-secretase compete for APP cleavage. The recombinant DNA molecules may be transfected, along with a reporter gene, into a cell line that processes APP into Aβ, and stable clones may be generated. Alternatively, the recombinant DNA molecules and reporter plasmid may be utilized in transient transfections.

Upon expression in cells, the APP/transcription factor fusion protein localizes to a non-nuclear membrane of the cell (e.g., the endoplasmic reticulum) due to the presence of the APP sequences in the fusion protein. In a manner similar to cleavage of APP, the fusion protein will then be cleaved, first by β-secretase and then by γ-secretase. γ-secretase cleavage releases the transcription factor from the membrane in which the APP/transcription factor fusion protein had been embedded, after which the transcription factor translocates to the nucleus and stimulates transcription of the reporter gene. Assuming no α-secretase cleavage, cleavage by both β-secretase and γ-secretase is required for release of the transcription factor and transactivation of the reporter gene in this assay since γ-secretase cleavage of APP is dependent on a short luminal domain, such as that generated by β- or α-secretase cleavage. Detection of a signal from the reporter gene product will thus serve as evidence of APP processing. In particular, since activation of the reporter gene requires both β-secretase and γ-secretase cleavage, the assay is capable of identifying inhibitors of both or either of these proteases.

Figure 27 is a schematic diagram depicting general features of the assay. The vertical bar in Figure 27A represents the fusion protein; the horizontal bar represents the non-nuclear membrane in which the fusion protein is embedded before processing. Figure 27B shows how the transcription factor portion of the fusion protein (with small amounts of the APP portion flanking it) has moved to the nucleus following release from the fusion protein by APP processing. In the nucleus, the

transcription factor is shown binding to a regulatory DNA sequence ("Transcription Factor Response Element") and activating transcription of the reporter gene.

The recombinant DNA molecules encoding the APP/transcription factor fusion protein and the reporter gene can be used to develop novel homogenous cell-based assays for the identification and assessment of inhibitors of APP processing which will be readily amenable to high throughput technology.

In one embodiment, the recombinant DNA molecules used in this invention comprise sequences encoding the amino terminal 651 amino acids of the 695 amino acid version of APP (Kang et al., 1987, Nature 325:733-736), including all the sequences necessary for the production of Aβ, as well as the C-terminal 32 amino acids of APP. The transcription factor is placed between the N-terminal and C-terminal portions of APP. The APP sequence may include a modification to increase the amount of β-secretase cleavage of the fusion protein. This modification involves mutating the K at position 612 of the α-secretase cleavage site to a V (K612V). Since α-secretase and β-secretase compete for APP cleavage, reducing or eliminating APP cleavage by α-secretase results in increased β-secretase cleavage, and allows the assay to detect β-secretase inhibitors more readily. In addition, the β-secretase cleavage site within APP (KM↓DA) (SEQ ID NO:34) may be modified, e.g., to that of a naturally occurring mutation (termed the "Swedish" mutation or NL↓DA) (SEQ ID NO:38) which has been shown to enhance β-secretase cleavage six-fold in cultured cells. Another possible modification is to replace the (KM↓DA) (SEQ ID NO:34) wild-type β-secretase cleavage site with the sequence (NF↓EV) (SEQ ID NO:40). The presence of NFEV in an amino acid sequence has been shown to enhance β-secretase cleavage by an even larger amount than the Swedish sequence. See U.S. Provisional Patent Application Serial No. 60/292,591 and U.S. Provisional Patent Application Serial No. 60/316,115, the disclosures of which are incorporated herein, in their entirety.

In a preferred embodiment, HIV-1 TAT exon I has been fused between sequences encoding the first 651 amino acids of APP_{695} and the last 32 amino acids of APP_{695} (APP-TAT-APPct32). Co-transfection of an expression vector comprising this construct with a reporter gene plasmid containing an HIV-1 LTR promoter that controls the transcription of a reporter gene leads to enhanced expression of the reporter gene. Other transcription factors that could be fused to APP_{1-651} include Gal4-VP16, the entire Gal4 protein, BIV TAT, HIV-2 TAT, SIV TAT, LexA-VP16, EBV Zta, Papillomavirus E2, or tissue or species specific homodimeric bHLH

transcription factors capable of activating transcription through specific DNA response elements, such as E12, E47, or Twist. The use of GAL4, BIV, HIV-2, or SIV TAT may be useful if it is desired to reduce the potency of the transactivator, thus reducing any background transactivation caused by non-specific cleavage of the fusion protein. To further reduce the potential for transactivation by TAT in the absence of β-secretase and γ-secretase cleavage, the TAT portion of the fusion protein may be altered to remove the N-terminal methionine and thus eliminate the possibility of aberrant translation of TAT through any potential internal ribosomal entry sites.

In some circumstances, high level expression of TAT has been found to be toxic to cells. Thus, when TAT is the transcription factor fused to APP in the methods of the present invention, it may be advantageous to utilize transient transfection with low amounts of the expression vector encoding the APP/TAT fusion protein. A set of preliminary experiments in which various amounts of the vector are transfected, in order to titrate acceptable levels of TAT, is recommended.

The reporter gene used will depend in large part upon the transcription factor fused to APP. The promoter used to drive the reporter gene will be LTR for TAT-based APP fusion proteins, or UAS (6x) for GAL4-VP16-based APP fusion proteins. In a particular embodiment, an LTR driving EGFP (enhanced green fluorescent protein, a brighter variant of GFP made by Aurora Biosciences, San Diego, CA) has been used to observe processing of an APP/TAT fusion protein. Under certain conditions, it may be desirable to use a less stable reporter, such as dsEGFP (a destabilized variant of EGFP made by Aurora Biosciences, San Diego, CA and marketed by Clontech, Palo Alto, CA) or a more potent reporter, such as β-lactamase. Alternatively, a stable HeLa cell line expressing LTR-β-galactosidase can be used. If the exquisite sensitivity of β-lactamase makes it less than optimal for a particular purpose, the LTR-β-galactosidase cell line may be exploited for this assay. Finally, under some circumstances Gal4-VP16 may prove to be optimal relative to TAT to reduce any inherent background problems associated with using the weakly but constitutively active LTR in the reporter plasmid, in which case the reporter plasmid could be UAS(6x)-β-lactamase (Aurora Biosciences, San Diego, CA).

A variety of cells are suitable for use in the methods of the present invention. Particularly preferred are eukaryotic, especially mammalian, cell lines. In particular embodiments, the cells are selected from the group consisting of: L cells L-M(TK$^-$) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), HEK293 (ATCC CRL

1573), HEK293T, Raji (ATCC CCL 86), CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BS-C-1 (ATCC CCL 26), T24 (ATCC HTB-4), PC12 cells, Jurkat cells, H4 cells (ATCC HTB-148), and MRC-5 (ATCC CCL 171).

To make the assay more amenable for ultra-high throughput screening, a non-adherent cell line, such as Jurkat, can be used.

Generally, the assays of the present invention employ cells that naturally express β-secretase and γ-secretase. However, it is possible to practice the invention in cells that lack the expression of one, or both, of these enzymes. In such cases, β-secretase and γ-secretase activity can be provided by the recombinant expression of these enzymes in the cells.

In one embodiment, the present invention provides a recombinant cell, preferably a eukaryotic cell, even more preferably a mammalian cell, and most preferably a human cell, where the cell expresses a fusion protein of APP and a transcription factor and the cell contains a reporter gene that can be activated by the transcription factor. The fusion protein comprises a portion of APP where that portion includes the regions of the β-secretase and γ-secretase cleavage sites fused to a transcription factor. The region of APP including the β-secretase and γ-secretase cleavage sites can be, e.g., a portion of APP that includes amino acids 589-651 of the 695 amino acid version of APP. This region is shown below.

↓ ↓ ↓

EEISEVKM DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV IA TVIVITLVMLKKK (SEQ ID NO:33)

↓

The β-secretase cleavage site is shown at position 596-597 (KM DA) (SEQ ID NO:34).

Two predominant cleavage sites of γ-secretase are shown at positions 636-637 and 638-639

↓ ↓

(GVV IA TV) (SEQ ID NO>35).

The fusion protein will be anchored in the membrane by the APP sequences shown above. The N-terminal portion of APP must include at least the β-

secretase cleavage site, and possibly several amino-acids N-terminal to the β-secretase cleavage site to make the assay sensitive to both β-secretase and γ-secretase inhibitors. In many cases, the APP sequences will include sequences further N-terminal than those shown above, including the signal sequence at the N-terminus of APP. In cases, where the APP signal sequence is not used, another signal sequence may be incorporated in the fusion protein. Such other signal sequences are known in the art.

In a related embodiment, the present invention provides a recombinant cell, preferably a eukaryotic cell, even more preferably a mammalian cell, and most preferably a human cell, where the above-described APP portion of the fusion protein contains a K612V mutation. The APP portion of this embodiment is shown below.

↓ ↓ ↓
EEISEVKM DAEFRHDSGYEVHHQVLVFFAEDVGSNKGAIIGLMVGGVV IA TVIVITLVMLKKK (SEQ ID NO:36)
↓

The β-secretase cleavage site is shown at position 596-597 (KM DA) (SEQ ID NO:34).

Two predominant cleavage sites of γ-secretase are shown at positions 636-637 and 638-639

↓ ↓
(GVV IA TV) (SEQ ID NO:35).

The underlined V at position 612 shows the change in sequence in the present invention from the wild-type K to the mutant V, which change provides for reduced cleavage by α-secretase.

In a related embodiment, the present invention provides a recombinant cell, preferably a eukaryotic cell, even more preferably a mammalian cell, and most preferably a human cell, where the above-described APP portion of the fusion protein contains the Swedish version of the β-secretase cleavage site as well as a K612V mutation. The APP portion of this embodiment is shown below.

↓ ↓ ↓
EEISEVNL DAEFRHDSGYEVHHQVLVFFAEDVGSNKGAIIGLMVGGVV IA TVIVITLVMLKKK (SEQ ID NO:37)
↓

The β-secretase cleavage site is shown at position 596-597 (NL DA) (SEQ ID NO:38).

Two predominant cleavage sites of γ-secretase are shown at positions 636-637 and 638-639

↓ ↓

(GVV IA TV) (SEQ ID NO:35).

The underlined V at position 612 shows the change in sequence in the present invention from the wild-type K to the mutant V, which change provides for reduced cleavage by α-secretase.

In a related embodiment, the present invention provides a recombinant cell, preferably a eukaryotic cell, even more preferably a mammalian cell, and most preferably a human cell, where the above-described APP portion of the fusion protein contains the NFEV version of the β-secretase cleavage site as well as a K612V mutation. The APP portion of this embodiment is shown below.

↓ ↓ ↓

EEISEVNF EVEFRHDSGYEVHHQVLVFFAEDVGSNKGAIIGLMVGGVV IA TVIVITLVMLKKK (SEQ ID NO:39)

↓

The β-secretase cleavage site is shown at position 596-597 (NF EV) (SEQ ID NO:40).

Two predominant cleavage sites of γ-secretase are shown at positions 636-637 and 638-639

↓ ↓

(GVV IA TV) (SEQ ID NO:35).

The underlined V at position 612 shows the change in sequence in the present invention from the wild-type K to the mutant V, which change provides for reduced cleavage by α-secretase.

The presence of both β-secretase and γ-secretase cleavage sites in the fusion proteins permits the assays of the present invention to detect inhibitors of both β-secretase and γ-secretase.

The recombinant host cells of the present invention can be further engineered to comprise a reporter gene construct. The reporter gene construct contains a reporter gene in operable linkage with a regulatory DNA sequence that confers on the reporter gene the property of being regulated by the transcription factor of the fusion protein. This regulation is such that expression of the reporter gene is low or absent without binding of the transcription factor to the regulatory DNA

sequence but, when the transcription factor is released from the fusion protein by APP processing, the transcription factor can move into the nucleus of the cell and bind to the regulatory DNA sequence, thereby activating transcription from the reporter gene.

Reporter genes desirably give rise to gene products which can be detected or quantitated, either in terms of amount of protein synthesized, enzymatic activity, fluorescence, luminescence, or some other phenotype. Suitable reporter gene products include firefly luciferase (de Wet et al., 1987, Mol. Cell. Biol. 7:725-737) or bacterial luciferase (Englebrecht et al., 1985, Science 227:1345-1347; Baldwin et al., 1984, Biochem. 23:3663-3667), β-lactamase, β-glucuronidase, β-galactosidase, green fluorescent proteins, enhanced green fluorescent protein, destabilized enhanced green fluorescent protein, red fluorescent protein, yellow fluorescent protein, cyan fluorescent protein, destabilized yellow fluorescent protein, destabilized cyan fluorescent protein, aequorin, chloramphenicol acetyl transferase (Alton & Vapnek, 1979, Nature 282:864-869), rat liver alkaline phosphatase (Toh et al., 1989, Eur. J. Biochem. 182:231-237), human placental secreted alkaline phosphatase (Cullen & Mallim, 1992, Meth. Enzymol. 216:362-368), and horseradish peroxidase, among others.

A preferred reporter gene is green fluorescent protein (GFP) or a modified GFP. Wild-type GFP has long been used in the art. Starting from green fluorescent protein, many modified versions have been derived with altered or enhanced spectral properties as compared with wild-type GFP. See, *e.g.*, U.S. Patent No. 5,625,048; International Patent Publication WO 97/28261; International Patent Publication WO 96/23810. Useful are the modified GFPs W1B and TOPAZ, available commercially from Aurora Biosciences Corp., San Diego, CA. W1B contains the following changes from the wild-type GFP sequence: F64L, S65T, Y66W, N146I, M153T, and V163A (see Table 1, page 519, of Tsien, 1998, Ann. Rev. Biochem. 67:509-544). TOPAZ contains the following changes from the wild-type GFP sequence: S65G, V68L, S72A, and T203Y (see Table 1, page 519, of Tsien, 1998, Ann. Rev. Biochem. 67:509-544). Wild-type nucleotide and amino acid sequences of GFP are shown in Figure 1 and SEQ ID NO: 1 of International Patent Publication WO 97/28261; in Figure 1 of Tsien, 1998, Ann. Rev. Biochem. 67:509-544; and in Prasher et al., 1992, Gene 111:229-233.

When expressing GFPs in mammalian cells, it may be advantageous to construct versions of the GFPs having altered codons that conform to those codons

preferred by mammalian cells (Zolotukhin et al., J. Virol. 1996, 70:4646-46754; Yang et al., 1996, Nucl. Acids Res. 24:4592-4593). Another way of improving GFP expression in mammalian cells is to provide an optimal ribosome binding site by the use of an additional codon immediately after the starting methionine (Crameri et al., 1996, Nature Biotechnology 14:315-319).

Transcription factors that are useful in the present invention are preferably those transcription factors that are not naturally expressed in the recombinant host cells. This is so the regulatory DNA sequence is not activated absent APP processing and release of the transcription factor from the fusion protein. Preferably, the transcription factor contains, or is engineered to contain, a nuclear localization signal. This is so that, after release, the transcription factor will move into the nucleus of the genetically modified host cells where it can bind to, and activate, the regulatory DNA sequence, leading to expression of the reporter gene. Transcription factors, as used in the present invention, do not include proteins that, after release from a fusion protein and translocation into the nucleus, repress transcription from a reporter gene.

Among the transcription factors that are useful in the present invention are: HIV1 TAT (in particular exon I of HIV1 TAT), Gal4-VP16, the entire Gal4 protein, BIV TAT, HIV-2 TAT, SIV TAT, LexA-VP16, EBV Zta, Papillomavirus E2, or one of the bHLH homodimeric transcription factors, E12, E47, or Twist.

Expression vectors are generally used to express the fusion protein in the recombinant cells. An expression vector contains recombinant nucleic acid encoding a polypeptide (e.g., an APP/transcription factor fusion protein) along with regulatory elements for proper transcription and processing. Generally, the regulatory elements that are present in an expression vector include a transcriptional promoter, a ribosome binding site, a transcriptional terminator, and a polyadenylation signal. Other elements may include an origin of replication for autonomous replication in a host cell, a selectable marker, a limited number of useful restriction enzyme sites, and a potential for high copy number.

A variety of expression vectors are known in the art and can be used in the present invention. Commercially available expression vectors which are suitable include, but are not limited to, pMC1neo (Stratagene), pSG5 (Stratagene), pcDNAI and pcDNAIamp, pcDNA3, pcDNA3.1, pCR3.1 (Invitrogen, San Diego, CA), EBO-pSV2-neo (ATCC 37593), pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12)

(ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pCI.neo (Promega), pTRE (Clontech, Palo Alto, CA), pV1Jneo, pIRESneo (Clontech, Palo Alto, CA), pCEP4 (Invitrogen, San Diego, CA), pSC11, and pSV2-dhfr (ATCC 37146). The choice of vector will depend upon the cell type in which it is desired to express the APP/transcription factor fusion protein, as well as on the level of expression desired, and the like.

The expression vectors can be used to transiently express or stably express the fusion protein. The transient expression or stable expression of transfected DNA is well known in the art. See, *e.g.*, Ausubel et al., 1995, "Introduction of DNA into mammalian cells," in Current Protocols in Molecular Biology, sections 9.5.1-9.5.6 (John Wiley & Sons, Inc.).

The recombinant host cells of the present invention are useful in methods of screening substances for the ability to inhibit APP processing. In one embodiment, the methods of the present invention comprise adding a candidate substance to a recombinant host cell comprising an APP/transcription factor fusion protein and a reporter gene and comparing the level of expression of the reporter gene protein in the presence and absence of the candidate substance, wherein the level of expression of the reporter gene protein is lower when the candidate substance inhibits processing of the APP/transcription factor fusion protein such that the transcription factor is not released, or is released in a lower amount, than in the absence of the substance.

The level of expression of the reporter gene protein is generally not measured directly. Rather, an indirect method is used. For example, fluorescence given off by the reporter gene protein may be detected or measured as, e.g., when the reporter gene product is a green fluorescent protein; or, some enzymatic activity of the reporter gene product may be detected or measured, e.g., when the reporter gene product is β-lactamase.

The candidate substance may be of any form suitable for entry into the cytoplasm of the recombinant cell or for contact with the cell's cytoplasmic membrane. Under appropriate conditions, the candidate substance may be allowed to freely diffuse into the cell, or the delivery of the substance may be facilitated by techniques and substances which enhance cell permeability, a wide variety of which are known in the art. Methods for increasing cell permeability include, without limitation, the use of organic solvents such as dimethylsulfoxide, liposomes,

application of electrical current, and physical means such as substance-coated teflon pellets.

The present invention provides a method of identifying a substance that inhibits APP processing comprising:

(a) providing a recombinant eukaryotic cell which:

(i) expresses a fusion protein comprising amino acids 589-651 of APP695 and a transcription factor where the transcription factor is fused in frame to the carboxyl terminus of amino acids 589-651 of APP695; and

(ii) comprises a reporter gene operably linked to a regulatory DNA sequence which is capable of being activated by the transcription factor;

(b) measuring the level of reporter gene product in the cell in the absence of the substance;

(c) adding the substance to the cell and measuring the level of reporter gene product in the cell in the presence of the substance;

where a decrease in the level of reporter gene product in the presence as compared to the absence of the substance indicates that the substance inhibits APP processing.

The manner in which the level of the reporter gene product is measured will be determined by the nature of the reporter gene and, often, the characteristics of the host cell. For example, if the reporter gene product itself is fluorescent, as for example, when a green fluorescent protein is the reporter gene product, fluorescence from the cell can be measured directly. When the reporter gene product has enzymatic activity, for example, when the reporter gene product is β-lactamase, known methods of measuring that enzymatic activity can be used.

For the sake of clarity, the above method is described in terms of "a" cell. In actual practice, the method will generally be carried on a large number of cells at one time. For example, the method will often be carried out in a well of a tissue culture plate, where, depending on the number of wells in the plate (and thus their size), there can be up to hundreds, thousands, or even several million cells. The step of "adding the substance to the cell" is generally carried out by simply adding the substance to the tissue culture medium in which the cells are present. After the substance is added to the cell, the cell and the substance are incubated for a period of time sufficient for the substance to inhibit APP processing, if the substance is actually

an inhibitor of APP processing. This period is usually from about 15 minutes to 48 hours, but may be somewhat more in unusual cases.

A convenient way of carrying out the method is to grow a population of the recombinant eukaryotic cells and then split the population into a portion that will be exposed to the substance and a portion that will not be exposed to the substance.

The recombinant eukaryotic cell is generally produced by transfection of an expression vector encoding the fusion protein and by transfection of a plasmid containing the reporter gene.

One skilled in the art would recognize that what is sought in terms of "a decrease in the level of reporter gene product in the presence as compared to the absence of the substance" is a non-trivial decrease. For example, if in the method described above there is found a 1% decrease, this would not indicate that the substance is an inhibitor of APP processing. Rather, one skilled in the art would attribute such a small decrease to normal experimental variance. What is looked for is a significant decrease. For the purposes of this invention, a significant decrease fulfills the usual requirements for a statistically valid measurement of a biological signal. For example, depending upon the details of the embodiment of the invention, a significant decrease might be a decrease of at least 10%, preferably at least 20%, more preferably at least 50%, and most preferably at least 90%.

In particular embodiments, amino acids 589-651 of APP_{695} contain a K612V mutation.

In particular embodiments, the cell is a mammalian cell. In particular embodiments, the cell is a human cell.

In particular embodiments, the method is used to screen a library of more than 1,000 substances. In other embodiments, the method is used to screen a library of more than 50,000 substances at a rate of more than 1,000 substances per 24 hours.

In particular embodiments, the fusion protein comprises a portion of APP that is selected from the group consisting of: amino acids 1-651 of APP_{695}, amino acids 50-651 of APP_{695}, amino acids 100-651 of APP_{695}, amino acids 150-651 of APP_{695}, amino acids 200-651 of APP_{695}, amino acids 250-651 of APP_{695}, amino acids 300-651 of APP_{695}, amino acids 350-651 of APP_{695}, amino acids 400-

651 of APP_{695}, amino acids 450-651 of APP_{695}, amino acids 500-651 of APP_{695}, and amino acids 550-651 of APP_{695}.

In related embodiments, the fusion protein does not comprise all of amino acids 589-651 of APP_{695}. Rather, the fusion protein comprises slightly fewer amino acids from APP. For example, the fusion protein might comprise slightly fewer amino acids of the β-secretase cleavage site: e.g., amino acids 590-651 of APP_{695}. Or the fusion protein might comprise slightly fewer amino acids of the γ-secretase cleavage site: amino acids 589-650 of APP_{695}; amino acids 589-649 of APP_{695}; amino acids 589-648 of APP_{695}; or amino acids 589-647. The fusion protein may even comprise slightly fewer amino acids from both ends, e.g., amino acids 590-647 of APP_{695}. What is important is that the portion of APP included in the fusion protein contains both the β-secretase cleavage site and the γ-secretase cleavage site.

In particular embodiments, the transcription factor is selected from the group consisting of: HIV-1 TAT, Gal4-VP16, the entire Gal4 protein, LexA-VP16, EBV Zta, Papillomavirus E2, one of the bHLH homodimeric transcription factors, including E12, E47, or Twist, or BIV TAT, HIV-2 TAT, or SIV TAT. A particular version of HIV-1 TAT suitable for use in the present invention is HIV-1 TAT exon I.

Fusion proteins suitable for use in the present invention can be selected from the group consisting of: APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:2); APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:4); APP(1-651)SW, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:6); APP(1-651)wt, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:8); APP(1-651)SW, TATexonI(M1L) APP (664-695) (SEQ ID NO:10); APP(1-651)wt, TATexonI(M1L) APP (664-695) (SEQ ID NO:12); APP(1-651)SW, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:14); APP(1-651)wt, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:16); APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:23); and APP(1-651)NFEV, K612V, GAL4-VP16(M1L) APP (664-695) (SEQ ID NO:25).

In some embodiments of the present invention, the amino acid sequences contributed to the fusion protein by the transcription factor constitute the carboxy terminal amino acid sequences of the fusion protein. In other embodiments, the transcription factor has other sequences fused to its carboxy terminus, as in the examples herein where amino acids 664-695 of APP_{695} are fused to the carboxy terminus of the transcription factor and therefore constitute the carboxy terminal

amino acid sequences of the fusion protein. Other portions of APP (e.g., amino acids 652-695 of APP695) could be used instead of amino acids 664-695 of APP695. In fact, it should be possible to extend the carboxy terminus of the transcription factor with almost any amino acid sequences, providing such sequences do not interfere with the ability of the transcription factor to move into the nucleus and activate transcription of the reporter gene once the transcription factor has been released from the fusion protein by the action of γ-secretase.

The present invention includes a method of identifying a substance that inhibits APP processing comprising:

(a) providing a recombinant eukaryotic cell which:

(i) expresses a fusion protein comprising an amino acid sequence from APP that is capable of being cleaved by both β-secretase and γ-secretase and a transcription factor where the transcription factor is fused in frame to the amino acid sequence from APP; and

(ii) comprises a reporter gene operably linked to a regulatory DNA sequence which capable of being activated by the transcription factor;

(b) measuring the level of reporter gene product in the cell in the absence of the substance;

(c) adding the compound to the cell and measuring the level of reporter gene product in the cell in the presence of the substance;

where a decrease in the level of reporter gene product in the presence as compared to the absence of the substance indicates that the substance inhibits APP processing.

In particular embodiments, the amino acid sequence from APP comprises:

589-651 of APP695;
590-651 of APP695;
589-650 of APP695;
590-650 of APP695;
589-649 of APP695;
590-649 of APP695;
589-648 of APP695;
590-648 of APP695;
589-647 of APP695; or

590-647 of APP_{695}.

In related embodiments, the amino acid sequence from APP contains the amino acid sequence NLDA (SEQ ID NO:38) at the β-secretase cleavage site instead of the wild-type sequence KMDA (SEQ ID NO:34).

In related embodiments, the amino acid sequence from APP contains the amino acid sequence NFEV (SEQ ID NO:40) at the β-secretase cleavage site instead of the wild-type sequence KMDA (SEQ ID NO:34).

The portion of the fusion protein that is derived from APP may contain mutations that are known in the art. Of particular interest are mutations that result in an increased proportion of Aβ being made in the form of Aβ1-42 rather than Aβ1-40. Such mutations are disclosed in the following publications (numbering is from the 770 amino acid version of APP):

Swedish (K670N, M671L): Mullan et al., 1992, Nature Genet. 1:345-347.

Flemish (A692G): Hendriks et al., 1992, Nature Genet. 1:218-221; Cras et al., 1998, Acta Neuropathol. (Berlin) 96:253-260.

Dutch (E693Q): Levy et al., 1990, Science 248:1124-1126.

Arctic (E693G): Nilsberth et al., 2001, Nature Neuroscience 4: 887-893.

Austrian (T714I): Kumar-Singh et al., 2000, Hum. Mol. Genet. 9:2589-2598.

French (V715M): Ancolio et al., 1999, Proc. Natl. Acad. Sci. (USA) 96:4119-4124.

Florida (I716V): Eckman et al., 1997, Hum. Mol. Genet. 6:2087-2089.

V717F: Murrell et al., 1991, Science 254:97-99.

V717G: Chartier-Harlin et al., 1991, Nature 353:844-846.

London (V717I): Goate et al., 1991, Nature 349:704-706.

L723P: Kwok et al., 2000, Ann. Neurol. 47:249-253.

I716F (also called I45F, referring to the position relative to the β-secretase cleavage site): This mutation in APP changes processing of Aβ almost exclusively to Aβ1-42. Lichtenthaler et al., 1999, Proc. Natl. Acad. Sci. (USA) 96:3053-3058.

As with many proteins, it may be possible to modify many of the amino acids of the fusion proteins described above and still retain substantially the same biological activity in terms of APP processing as for the original fusion protein. Thus, the present invention includes modified fusion proteins which have amino acid deletions, additions, or substitutions but that still retain substantially the same properties with respect to APP processing as the fusion proteins described herein. It is generally accepted that single amino acid substitutions do not usually alter the

biological activity of a protein (see, *e.g.*, Molecular Biology of the Gene, Watson *et al.*, 1987, Fourth Ed., The Benjamin/Cummings Publishing Co., Inc., page 226; and Cunningham & Wells, 1989, Science 244:1081-1085). Accordingly, the present invention includes fusion proteins where one amino acid substitution has been made in the fusion proteins described herein where the fusion proteins still retain substantially the same properties with respect to APP processing as the fusion proteins described herein. The present invention also includes fusion proteins where two or more amino acid substitutions have been made in the fusion proteins described herein where the fusion proteins still retain substantially the same properties with respect to APP processing as the fusion proteins described herein. In particular, the present invention includes embodiments where the substitutions are conservative substitutions.

With the exception of Figure 18, the nucleotide and amino acid sequences of APP disclosed herein contain a minor difference compared to APP sequences that are usually reported in the literature. For the sequences disclosed herein with such a difference, the nucleotide at position 367 is an A rather than a G, as in most published APP sequences. This change results in a conservative substitution in the corresponding APP amino acid sequence. Thus, the amino acid sequences disclosed herein with such a difference have an I rather than a V at position 123. This difference does not affect the properties of the fusion proteins for the purposes of the present invention. Therefore, fusion proteins having the APP sequence reported in the literature with an G at nucleotide position 367 and a V at amino acid position 123 and the fusion proteins disclosed herein with an A at nucleotide position 367 and an I at amino acid position 123 are to be considered equivalents for the purposes of the present invention.

The Gal-VP16 sequences disclosed herein contain two changes from the usual published sequences. There is T to C change at nucleotide position 2131 that causes a S to P change at amino acid position 712; there is A to C change at nucleotide position 2301 that does not change the amino acid sequence. It is expected that Gal-VP16 proteins containing the usual sequences reported in the literature will also be suitable for use in the present invention.

The methods of the present invention can be used to screen libraries of substances or other sources of substances to identify substances that are inhibitors of β-secretase or γ-secretase. Such identified inhibitory substances can serve as "leads"

for the development of pharmaceuticals that can be used to treat patients having Alzheimer's disease or in a prophylactic manner to prevent or delay the development of Alzheimer's disease. Such leads can be further developed into pharmaceuticals by, for example, subjecting the leads to sequential modifications, molecular modeling, and other routine procedures employed in the pharmaceutical industry. The inhibitors of APP processing identified by the present invention may also be tested in animal models of Alzheimer's disease such as the various transgenic mouse models that are known in the art.

Although a wide variety of substances can be screened by the methods of the present invention, preferred substances for screening are libraries of small molecule compounds. Small molecule compounds are preferred because they are more readily absorbed after oral administration, have fewer potential antigenic determinants, and are more likely to cross the blood/brain barrier than larger molecules such as nucleic acids or proteins.

Once identified by the methods of the present invention, the candidate small molecule compounds may then be produced in quantities sufficient for pharmaceutical testing and formulated in a pharmaceutically acceptable carrier (see, e.g., Remington's Pharmaceutical Sciences, Gennaro, A., ed., Mack Publishing, 1990, for suitable methods). The candidate compounds may be administered to cell lines relevant to Alzheimer's disease, animal models of Alzheimer's disease, or Alzheimer's disease patients.

The numbering of the amino acids in APP used herein is based on the 695 amino acid version of APP described in Kang et al., 1987, Nature 325:733-736. There are two other major versions of APP, having 751 amino acids and 770 amino acids (see, Ponte et al., 1988, Nature 331:525-527 for the 751 amino acid version and Kitaguchi et al., 1988, Nature 331:530-532 for the 770 amino acid version). One skilled in the art will understand how to translate the numbering used herein, based on the 695 amino acid version of APP, into the corresponding numbering for other versions of APP. For example, some of the APP/transcription factor fusion proteins of the present invention contain the K612V mutation, based on the numbering of the 695 amino acid version. This mutation would correspond to a K668V mutation in the 751 amino acid version and a K687V mutation in the 770 amino acid version.

Therefore, when a "K612V" mutation is referred to herein, it will be understood that such reference also includes a K668V mutation of the 751 amino acid version of APP as well as a K687V mutation of the 770 amino acid version of APP.

Similarly, the portion of APP referred to as APP_{1-651} herein, based on the 695 amino acid version, will be understood to mean also APP_{1-707} of the 751 amino acid version and APP_{1-726} of the 770 amino acid version.

If desired, inhibitors that are identified by the methods of the present invention can be further tested to determine which step in APP processing they affect. Assays that are known to be specific for the various steps of APP processing can be used for this purpose. For example, the assay of Karlström et al., (Journal of Biological Chemistry papers in press, published on December 13, 2001 as Manuscript C100649200) is only capable of detecting inhibitors of γ-secretase and cannot also detect inhibitors of other steps of APP processing such as, e.g., inhibitors of β-secretase. If an inhibitor identified by the methods of the present invention is found to also be an inhibitor when tested in the assay of Karlström et al., then that inhibitor is at least a γ-secretase inhibitor. It is still possible that that inhibitor could inhibit other steps in APP processing as well. Further tests known in the art can determine this.

The present invention may be modified so as to provide methods of determining at which step of APP processing a known inhibitor of APP processing exerts its effect. The known inhibitor may be one that has been identified by the methods of the present invention or by some other method. The modification to the present invention consists in mutating the β-secretase site in a fusion protein so that β-secretase cleavage can no longer occur at the site or occurs at a very much reduced level. Providing that the fusion protein contains a cleavable α-secretase site, the fusion protein can still be used in the methods of the present invention. However, this fusion protein (with a mutated β-secretase site) can no longer detect β-secretase inhibitors. Therefore, if the known APP processing inhibitor still functions as an APP processing inhibitor in this modified version of the invention, then the known inhibitor cannot be a β-secretase site inhibitor but instead must exert its effect downstream of β-secretase.

Suitable mutations of the β-secretase site include the following. All the sequences are for amino acid positions 594-598 of APP_{695}.

VNFAV (SEQ ID NO:41): This mutation shows decreased β-secretase cleavage relative to the wild type, KMDA (SEQ ID NO:34), sequence.

VKVDA (SEQ ID NO:42): Vassar et al., 1999, Science 286:735-741. This mutant was tested in vitro only, but purified β-secretase failed to cleave a 30-amino acid peptide containing this sequence.

WKMDA (SEQ ID NO:43), VKADA (SEQ ID NO:44), VKKDA (SEQ ID NO:45), VKEDA (SEQ ID NO:46), VKIDA (SEQ ID NO:47), VKMIA (SEQ ID NO:48), VKMNA (SEQ ID NO:49), VKMEA (SEQ ID NO:50), VKMDE (SEQ ID NO:51), VKMDK (SEQ ID NO:52): Citron et al., 1995. Neuron 14:661-670. These mutations decreased Aβ production 4-20X relative to p3 production in cultured cells.

Fusion proteins can be constructed by use of the polymerase chain reaction (PCR) to amplify desired portions of APP and transcription factors, which can be then be cloned into expression vectors by methods well known in the art. Primers for PCR will generally include a small part of the APP or transcription factor as well as convenient cloning sites and/or linker peptide sequences. The PCR primers can be used to amplify the desired APP or transcription factor fragments from sources such as previously cloned APP or transcription factors, cDNA libraries, or genomic libraries. The amplified APP and transcription factor sequences can be cloned into suitable expression vectors. Methods of PCR and cloning are well known in the art and can be found in standard reference materials such as those listed below.

Standard techniques for cloning, DNA isolation, amplification and purification, for enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like, and various separation techniques are known and commonly employed by those skilled in the art. A number of standard techniques are described in Sambrook et al. (1989) Molecular Cloning, Second Edition, Cold Spring Harbor Laboratory, Plainview, N.Y.; Maniatis et al. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, Plainview, N. Y.; Wu (ed.) (1993) Meth. Enzymol. 218, Part I; Wu (ed.) (1979) Meth. Enzymol. 68; Wu et al. (eds.) (1983) Meth. Enzymol. 100 and 101; Grossman and Moldave (eds.) Meth. Enzymol. 65; Miller (ed.) (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.; Old and Primrose (1981) Principles of Gene Manipulation, University of California Press, Berkeley; Schleif and Wensink (1982) Practical Methods in Molecular Biology; Glover (ed.) (1985) DNA Cloning Vol. I and II, IRL Press, Oxford, UK; Hames and Higgins (eds.) (1985) Nucleic Acid Hybridization, IRL Press, Oxford, UK; Setlow and Hollaender (1979) Genetic Engineering: Principles

and Methods, Vols. 1-4, Plenum Press, New York, and Ausubel et al. (1992) Current Protocols in Molecular Biology, Greene/Wiley, New York, N.Y..

PCR reactions can be carried out with a variety of thermostable enzymes including but not limited to AmpliTaq, AmpliTaq Gold, or Vent polymerase. For AmpliTaq, reactions can be carried out in 10 mM Tris-Cl, pH 8.3, 2.0 mM $MgCl_2$, 200 μM of each dNTP, 50 mM KCl, 0.2 μM of each primer, 10 ng of DNA template, 0.05 units/μl of AmpliTaq. The reactions are heated at 95°C for 3 minutes and then cycled 35 times using suitable cycling parameters, including, but not limited to, 95°C, 20 seconds, 62°C, 20 seconds, 72°C, 3 minutes. In addition to these conditions, a variety of suitable PCR protocols can be found in PCR Primer, A Laboratory Manual, edited by C.W. Dieffenbach and G.S. Dveksler, 1995, Cold Spring Harbor Laboratory Press; or PCR Protocols: A Guide to Methods and Applications, Michael et al., eds., 1990, Academic Press.

It is desirable to sequence the DNA encoding the fusion proteins, or at least the junction regions of the various portions (APP, transcription factor, linkers) of the fusion protein in order to verify that the desired portions have in fact been obtained, joined properly, and that no unexpected changes have been introduced into the sequences by the PCR reactions.

Suitable PCR primers for amplification of DNA sequences for use in the present invention can be readily designed by those of skill in the art. Examples of such primers are shown below.

5'-GGA GAG GAT ATC ATG GAG CCA GTA GAT CC-3' (SEQ ID NO:53) can be used to amplify the 5' portion of HIV-1 TAT exon I.

5'-TAC ATG GCG GCC GCC TAC TTA CTG CTT TG-3' (SEQ ID NO:54) can be used to amplify the 3' portion of HIV-1 TAT exon I.

5'-GGA TGT GAT ATC TTT CTT CTT CAG CAT CAC CAA GG-3' (SEQ ID NO:55) can be used to amplify the 3' portion of DNA encoding amino acids 1-651 of APP, i.e., the transmembrane region of APP.

The following non-limiting examples are presented to better illustrate the invention.

EXAMPLE 1

Transfection of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI with pMM321

The following example demonstrated that an APP/TAT fusion construct will transactivate a reporter gene in which the HIV1 LTR regulatory DNA sequence controls the expression of enhanced green fluorescent protein (EGFP). The following also serves as an example of the kind of preliminary routine variations of fusion protein levels and inhibitor levels that may be advantageous to test in the practice of the present invention. Such routine variations are often helpful in validating the assays before a large scale screening project is undertaken.

The APP/TAT fusion construct is referred to as "pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI" (see Figure 26) and contains the HIV1 TAT exon 1 fused just after the transmembrane domain of APP. This construct is also shown in outline form in Figure 1B. "pMM321" refers to a reporter gene plasmid consisting of the HIV1 LTR driving the transcription of enhanced green fluorescent protein (see Figure 25). As a positive control for TAT expression, a construct in which TAT was under the control of a strong, constitutive promoter (referred to as "pUCd5TAT"; see Figure 24) was used.

METHODS:

1. Day 1: Pass HEK 293T cells into 2 x 6 well dishes at 1 x 10^5 cells/well.

2. Day 2: Transfect cells with 9 µL Fugene and 0.125 µg pMM321 and various amounts of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI.

Plate 1:
1. 5 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
2. 2.5 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
3. 1.25 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI

4. 0.625 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
5. 0.312 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
6. 0.156 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI

Plate 2:
1. 0.08 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
2. 0.04 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
3. no pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
4. 0.625 pUCd5TAT
5. 0.625 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI and 1 μg pMM321
6. 1 μg pMM321

Six hours post-transfection, green cells were only observed in plate 2, #5.

3. Day 3: The fluorescence intensity of the transfected cells was observed and recorded.

4. Day 4: The fluorescence intensity of transfected cells was observed and recorded.

RESULTS:

Co-transfection with pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI increased GFP expression in the cells.

Day 3:

- 5 μg – no green cells
- 2.5 μg – no green cells (too much DNA for these two transfections?)
- 1.25 μg - many bright and dim green cells (see photographs and figure in ancillary data)
- 0.625 μg - bright and dim green cells but fewer than at 1.25 μg
- 0.312 μg - no difference obvious between 0.625 μg and 0.312 μg

- 0.156 μg - very few green cells
- 0.08 μg - very few green cells
- 0.04 μg - very few green cells
- no pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI - extremely few (if any) green cells
- 0.625 μg pUCd5TAT - cells were extremely bright, not necessarily more in number than with pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
- 0.625 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 1 μg pMM321 - many bright and dim green cells.
- 1 μg pMM321 many-fold fewer green cells, some bright, most dim.

Day 3 - changed media (saved 1 mL conditioned media from wells Plate 1- 3, 4, 5, 6; Plate 2 - 1, 2, 3, 5, 6). Added fresh media with 10 μM of L-685,458 (a potent, cell permeable γ- secretase inhibitor) to wells Plate 1 - 3, 4, 5, 6; Plate 2 - 3, 4, 5, 6. Waited 48 hours to observe loss of fluorescence since GFP is so stable.

After 48 hours, all wells appeared brighter than at 24 hour time point. This does not necessarily mean that the inhibitor was ineffective, or that the assay did not work, since there were no controls run where the inhibitor was not added. However, it does suggest that under these conditions it may be preferable to add the inhibitor at the time of transfection to shut down γ-secretase as soon as possible and avoid release of TAT and induction of GFP.

EXAMPLE 2

Transfection of APP(1-651)SW, K612V-(M1L)TATexonI into HEK293T and H4 cells accompanied by inhibition of γ-secretase activity with L-685,458

The following example demonstrates the operation of the invention in HEK293T cells and H4 cells and shows inhibition of APP processing (and thus TAT release) by treatment with a known γ-secretase inhibitor. "pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI," "pMM321," and "pUCd5TAT" are the same as in Example 1. H4 cells (ATCC HTB-148) are a neuronal cell line.

METHODS:

Day1: Plated out 2 x 6 well plates of HEK293T cells and 2 x 6 well plates of H4 cells at $1x10^5$ cells/well.

Day 2: Transfected cells with 2 μg total DNA - pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI and carrier (a pET-IN plasmid).

Plate1:

1,2: 1 μg pMM321 + 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI

3,4: 1 μg pMM321 + 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 0.9 μg carrier

5: 1 μg pMM321 + 1 μg carrier (added too much carrier to this well in H4 cells)

6: 1 μg pMM321 + 0.1 μg pUCd5TAT + 0.9 μg carrier

Plate 2:

1,2: 0.1 μg pMM321 + 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 0.9 μg carrier

3,4: 0.1 μg pMM321 + 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 1.8 μg carrier (added too 0.5X carrier to this mix in 293T cells)

5: 0.1 μg pMM321 + 1.9 μg carrier

6. 0.1 μg pMM321 + 0.1 μg pUCd5TAT + 1.8 μg carrier

Transfections for HEK293T cells: 9 μL Fugene/well. Combined with DNA in Optimem and incubated and added to cells according to manufacturer's instructions.

Transfections for H4 cells: 6 μL Fugene/well. Combined with DNA in Optimem and incubated and added to cells according to manufacturer's instructions.

Added 10 μM L-685,458 to Plates 1 and 2, wells 2 and 4 for both cell types within 1 hour of transfection. Observed cells periodically.

Took pictures at 24, 46 hours after transfection, using AE lock to keep exposures constant between wells.

RESULTS:

Both H4 and 293T cells turned much brighter green in the presence of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI

At 24 hours:

H4 cells:

Plate 1: 1 ug pMM321

1. 1 ug pMM321 + 1 ug pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: Many bright and dim green cells (good transfection efficiency as well)
2. 1 ug pMM321 + 1 ug pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 μM L-685,458: Also many bright and dim green cells, but reduced compared with well #1
3. Very few green cells (a few per field)
4. Very few green cells
5. A few dimly green cells
6. Some induction with 0.1 μg pUCd5TAT but still relatively few cells.

Plate 2: 0.1 μg pMM321

1. 0.1 μg pMM321 + 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: ~10 bright green cells/field and the rest are dim green
2. 0.1 μg pMM321 + 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 μM L-685,458: ~3-5 bright green cells/field, some dim green, and some not green.
3. No green cells
4. No green cells
5. No green cells
6. A few bright green cells

HEK293T cells:

Plate 1: 1 μg pMM321

1. + 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: of 15 cells: 6 dim, 4
medium, 5 bright
2. + 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 μM L-685,458:
of 15 cells: 6 very dim, 5 dim, 1 medium, 3 bright
3. +0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: many more green
cells than 1 μg, lots of strong, bright green cells
4. + 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 μM L-685,458:
Fewer bright green cells/field but intensity does not appear strongly diminished
5. 1 μg pMM321 alone: Most cells in the field expressing dim to medium levels of GFP
6. enhancement by 0.1 μg pUCd5TAT

Plate 2: 0.1 μg pMM321
1. + 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: Bright and medium
green cells
2. + 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 μM L-685,458:
Bright, medium, and dim green cells
3. + 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: Bright, medium,
and dim green cells
4. + 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 μM L-685,458:
Bright, medium, and dim green cells
5. 0.1 μg pMM321 alone: Most expressing cells have dim GFP, a few medium to bright
cells
6. Enhancement by 0.1 μg pUCd5TAT

Changed media on cells at 24 hours past transfection. Kept 10 μM L-685,458 on cells in wells 2 and 4.

At 46 hours after transfection, examined the wells again. Lots of floating cells in all wells, all cell types. Highest number of floaters in 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI lanes.

Took some photographs under fluorescent and white light (white light at low intensity) to reveal fluorescent and non-fluorescent cells. Conducted a subjective analysis of the photographs to see if the amount of inhibition by 10 μM L-685,458 was in any way quantifiable. Counted bright (white in the middle), strong (blue middle), medium (green) and dim/non-fluorescent cells and determined the approximate fraction of each level of expression. Results follow:

TABLE 1

293T cells transfected with X μg pMM321 (first number in left column) and X μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI (second number in left column)

Transfection	# Bright	# Strong	# Med	# Non	% Bright	%Strong	% Med	%Non
1 μg + 1μg	none	8	81	316		2	20	78
1 + 1 + cmpd	none	5	66	521		0.8	11	88
1 + 0.1	41	77	143	61	12	24	44	19
1+ 0.1+cmpd	23	32	149	194				

The results shown in Table 1 indicate that the presence of L-685,458 ("cmpd") caused fewer strong and medium fluorescing cells as well as more non-fluorescent cells in the first run; in the second run, L-685,458 caused fewer bright and strong fluorescing cells as well as more non-fluorescent cells (with slightly more medium fluorescing cells). Overall, these data clearly indicate that the presence of an inhibitor of APP processing such as L-685,458 can be identified by the present invention.

1 mL of conditioned media from each well was analyzed for production of Aβ. Higher than background levels of Aβ were observed in 293T cells after transfection

with 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI and 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI and higher than background levels of Aβ in H4 cells after transfection with 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI, but not 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI (no enhancement of GFP was observed with 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI either). Aβ was completely inhibited to background levels by 10 μM L-685,458. Surprisingly, substantial inhibition of GFP was not observed with 10 μM L-685,458.

100,000 cells from each well were trypsinized and placed in 0.1 mL phenol red-free media in a Costar 96-well dish and read using the fluorometer. The results are shown below:

TABLE 2

	1	2	3	4	5	6	7	8	9	10	11	12
A			1ugAPPTAT	10uM 458	0.1ugAPPtat	10 uM458	no tat	0.1ug pUCd5TAT				
B	293T	1ug LTRGFP	17670	14321	65535	65535	14890	65535				
C		0.1ug LTRGFP	9976	10491	17677	14790	9624	25735				
D	H4	1ug LTRGFP	21307	25307	7175	7136	7147	7277				
E		0.1ug LTRGFP	9574	10031	7317	6957	6946	7247				
F		Blank	7498	7124	7570	7454	5774	7638				
G												
H												

In Table 2, "APPtat" is pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI. "458" is L-685,458. "LTRGFP" is pMM321.

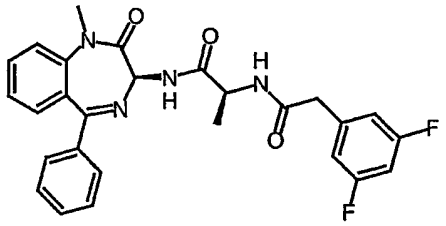
0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI with and without compound exceeded the maximum reading of the fluorometer, as did the addition of 0.1 μg pUCd5TAT to cells transfected with 1 μg pMM321.

1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 1 μg pMM321 incrementally increased the amount of fluorescence relative to 1 μg pMM321 alone, and this was reduced to background levels by 10 μM L-685,458. Inhibition of fluorescence was also observed in 293T cells transfected with 0.1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 0.1 μg pMM321. No inhibition of fluorescence was observed in H4 cells under any transfection conditions.

EXAMPLE 3

Use of APP(1-651)SW, K612V-TATexonI in H4 cells

L-875,532 is a known γ-secretase inhibitor having the structure shown below. It is described and details of its synthesis are disclosed in Seiffert et al., 2000, J. Biol. Chem. 275:34086-34091.



L-875532

Compound X is a β-secretase inhibitor.

pRBR186 (Figure 22A) is an expression vector containing DNA sequences encoding full-length APP containing the Swedish mutation and the K612V mutation. pRBR186 does not contain a transcription factor fused to the APP sequences.

pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI is an expression vector that directs the expression of a the fusion protein APP(1-651)SW, K612V-(M1L)TATexonI in mammalian cells. This fusion protein contains the first 651 amino acids of APP (with a Swedish version of the β-secretase cleavage site as well as the K612V mutation) fused in frame to exon I of HIV1 TAT, which has been modified with a Met1-Leu mutation. A schematic diagram of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI is shown in Figure 26A. The nucleotide sequence of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI (SEQ ID NO:22) is shown in Figure 26B-G.

METHODS:

1. H4 cells (ATCC HTB-148) were transfected with the various constructs listed below using 6 μL Fugene per 100 μL Optimem and 100 μL Optimem per well (6-well dishes). Transfection reactions were incubated for 30 minutes prior to adding 100 μL dropwise onto wells.

Transfections were done as follows:

1. 1 μg pMM321 (Figure 25A-D) and 1 μg pcDNA3.1 backbone
1a. 1 μg pMM321 and 1 μg pcDNA3.1 (Invitrogen, San Diego, CA) backbone. Prior to transfection, 10 μM L-875,532 (γ-secretase inhibitor) was added to the well.
2. 1 μg pMM321 and 1 μg pRBR186 (Figure 22A; APP expression vector; processing and inhibition of processing control)
2a. 1 μg pMM321 and 1 μg pRBR186. Prior to transfection, 10 μM L-875,532 was added to cells

(transfection solution for 3-5 were prepared in bulk)

3. and 3a. 1 μg pMM321 and 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
4. and 4a. 1μg pMM321 and 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI. Prior to transfection, 10 μM L-875,532 was added to the two wells.
5. and 5a. 1 μg pMM321 and 1 μg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-

(M1L)TATexonI. Prior to transfection, 10 μM Compound X was added to the two wells

6. 1 μg pMM321 and 1 μg pUCd5TAT (Figure 24).
7. 1 μg pMM321 and 1 μg pUCd5 TAT. Prior to transfection, 10 μM L-875,532 was added to the cells.

RESULTS:

Cells were assessed by eye under a fluorescence microscope the morning following transfection (~20 hrs).

1 and 1a, 2 and 2a. Weak fluorescence

3 and 3a. Much stronger fluorescence

4 and 4a. Clear inhibition of fluorescence

5 and 5a. Possible inhibition of fluorescence, but doesn't look that great

6 and 7. Almost blindingly fluorescent.

At approximately 48 hours, cells were trypsinized, spun down, and resuspended in 100 μL PBS. The cellular contents of each well of the transfection plates were placed into one well of a 96-well fluorescence plate. Fluorescence was analyzed using the FLUOstar (485 excitation/538 emission). The results are shown in Table 3.

TABLE 3

Transient transfections in H4 cells	Fluor Units
pMM321	4484
pMM321 + L-875,532	3443
pMM321 + pRBR186	2735
pMM321 + pRBR186 + L-875,532	2161
pMM321 + APP-TAT-ct32	20177
pMM321 + APP-TAT-ct32 + L-875,532	8283
pMM321 + APP-TAT-ct32 + Compound X	11946
pMM321 + pucd5-TAT	61102

In Table 3, "APP-TAT-ct32" refers to pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI.

For a graphical presentation of these results, see Figure 19. In Figure 19, "LTR-GFP" refers to pMM321; "APP-TAT-ct32" refers to pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI. Compare the bar labeled "LTR-GFP + APP-TAT-ct32" with the bars labeled "LTR-GFP + APP-TAT-ct32 + L-875,532" and "LTR-GFP + APP-TAT-ct32 + Compound X." Inhibition by both the β-secretase inhibitor (Compound X) and the γ-secretase inhibitor (L-875,532) is easily identified by the present invention.

Conclusions:

- The data indicate that the expression of the fusion protein APP(1-651)SW, K612V-(M1L)TATexonI enhances transactivation through the LTR of pMM321 in a manner that depends on APP processing.
- APP(1-651)SW, K612V-(M1L)TATexonI expressing cells were 6X brighter than pMM321 cells alone.
- Treatment with L-875,532 decreased fluorescence 2.5X.
- Treatment with Compound X decreased fluorescence 1.7X.
- Expression of TAT via pucd5-TAT was almost blinding and was 19X above pMM321 alone, indicating that APP(1-651)SW, K612V-(M1L)TATexonI expression did not lead to levels of GFP as high as TAT alone. Despite the decreased activation shown by TAT when provided by the fusion protein, as compared with TAT driven by the AMLP (adenovirus major late promoter) in pucd5-TAT, the assay was easily able to identify both the β-secretase and the γ-secretase inhibitors.
- Control plasmids (pMM321 and pMM321 + pRBR186) were dimly fluorescent and were not inhibited by L-875,532.

A lower level of inhibition by the β-secretase inhibitor is to be expected since the K612V mutation decreases alpha-secretase activity by 95% and thus some alpha-secretase cleavage is to be expected.

EXAMPLE 4

Construction of pcDNA3.1 (+) zeo APP(1-651)SW, K612V, GAL4-VP16(M1L) APP (664-695)

1. The GAL4-VP16 insert was prepared by PCR from pCR2.1 GAL4VP16 (Figure 20) (Invitrogen, San Diego, CA). The PCR was performed to eliminate the N-terminal methionine by changing this methionine into a leucine.

40 ng pCR2.1 GAL4-VP16
0.2 µL GAL4VP16 5' oligo at 250 µM:
5'-CTGAGATATCAAGCTACTGTCTTCTATCGAACAAGC-3' (SEQ ID NO:56)
EcoRV site underlined
0.2 µL GAL4VP16 3' oligo (at 250 µM): 5'-
GCGCGATATCCCCACCGTACTCGTCAATTCC-3' (SEQ ID NO:57)
EcoRV site underlined
5 µL 10X Buffer
8 µL 25 mM $MgCl_2$
4 µL PCR dNTPs
0.25 µL AmpliTaq Gold
27.35 µL water

Cycle:

Purified reactions using a Qiaquick column
Digested entire reaction using EcoRV
Ran the DNA on a 1% gel. Excised the band and purified using a QiaQuick gel purification kit

2. Digested pcDNA3.1 APP(1-651)/APP(664-695) with EcoRV and SAP treated. pcDNA3.1 APP(1-651)/APP(664-695) is an intermediate plasmid formed in the procedure described in Example 6. pcDNA3.1 APP(1-651)/APP(664-695) the first 651 amino acids of APP (with a Swedish version of the β-secretase cleavage site as well as the K612V mutation) fused in frame to the last 32 amino acids of APP.

3. Ligated pcDNA3.1 APP(1-651)/APP(664-695) –EcoRV digested to GAL4VP16 (EcoRV digested)

4. At this point, it was realized that the 3' PCR primer for GAL4-VP16 put the APP(664-695) fragment out of frame. The APP(664-695) fragment was then re-PCR'd using the following protocol:

1 µL pcDNA3.1 APP(1-651)-Gal4VP16-APP(664-695)
50 nM APP NotI 5'ct32 in frame with GAL4-VP16
50 nM APP Noti 3' ct32 (5' (p)CTGCTGTGGCGGCCGCCTAGTTCTGCATCTGCTC) (SEQ ID NO:58)
NotI site underlined
1 µL PCR dNTPs (10 mM each dNTP, Roche)
5 µL 10X Expand Buffer with $MgCl_2$
40 µL water
1 µL Expand polymerase (Roche)

The PCR fragment was run on a 4% agarose gel and gel-purified using a QiaQuick gel purification column
The fragment was digested with NotI and purified using a QiaQuick PCR purification column.

5. APP(1-651)-Gal4VP16-APP(664-695) was re-miniprepped. Miniprep #1 was digested with NotI, run on a 1% gel, the upper band was then isolated and SAP-treated.

6. APP(1-651)-Gal4VP16/NotI digested/SAP-treated was ligated to APP(664-695).

7. Minipreps containing inserts were sequenced to verify the orientation of the insert.

EXAMPLE 5

Construction of pcDNA3.1 zeo (+) APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695)

This procedure replaced a fragment of APP in pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) that contained the Swedish mutation with a corresponding fragment from pRBR121 containing the wild-type β-secretase cleavage site rather than the Swedish β-secretase cleavage site.

1. pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) (Figure 21B) was digested with SnaBI and then EcoRI

17 µL miniprep DNA
2 µL 10X buffer
1 µL SnaBI (NEB)

Digest was purified using Qiaquick PCR purification kit. Entire digest was then cleaved with EcoRI for 2 hours.

2. pRBR121 (Figure 21A) was digested with SnaBI and then EcoRI

5 µg pRBR121
5 µL 10X buffer
2.5 µL SnaBI (NEB)
q.s. 50 µL with water

The digest was purified using a Qiaquick PCR purification kit. The entire digest was then cleaved with EcoRI for 2 hours.

3. Both digests were run out on a 1% agarose gel. From the pRBR121 lane, the 2.4 kb SnaBI-EcoRI fragment containing the wild-type β-secretase cleavage site was isolated.

From pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) digest, BOTH the 5 kb SnaBI-EcoRI backbone fragment AND the 200 bp EcoRI-EcoRI fragment were isolated (see Figure 21B).

4. A three-part ligation using equal molar ratios of the three fragments was carried out:
The assumption was made that, since the starting plasmids were of similar sizes and the same amount was digested for each plasmid, the recovered fragments would be in approximately equal molar ratios.

Vector alone:
1μL APP-TAT-ct32 SnaBI/EcoRI 5Kb fragment
7 μL water
2 μL 5X buffer
10 μL 2X buffer (Roche rapid ligation kit)
1 μL T4 ligase

1+1+1
1 μL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) SnaBI/EcoRI 5KB fragment
1 μL pRBR121 SnaBI/EcoRI 2.4 Kb insert
1 μL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) EcoRI/EcoRI 200 bp insert
5 μL water
2 μL 5X buffer
10μL 2X buffer
1 μL T4 ligase

1+1+...1 (in this 3-pt ligation, the ligation of two of the fragments together was done 1st, then the third fragment was added)
1 μL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) SnaBI/EcoRI 5Kb backbone
1 μL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) EcoRI/EcoRI 200 bp insert

5 µL water
2 µL 5X buffer
10 µL 2X buffer
1 µL T4 ligase
waited 5 minutes
then added 1µL pRBR121 SnaBI/EcoRI 2.4 kb insert

1+1+3
1 µL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) SnaBI/EcoRI 5 kb backbone
1 µL pRBR 121 SnaBI/EcoRI 2.4 kb insert
3 µL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) EcoRI/EcoRI 200 bp insert
3 µL water
2 µL 5X buffer
10 µL 2X buffer
1 µL T4 ligase

Transformed and plated out 200 µL. The number of colonies in the vector + insert ligations far exceeded the number of colonies in the vector alone ligation. Picked 12 colonies from 1+ 1+...1.
Picked 6 colonies from 1+3.
Miniprepped
Digested with EcoRI to ensure that small 200 bp fragment was incorporated.
RESULTS
Minipreps #10 and 15 contained 200 bp EcoRI fragment.

Oriented with Bam HI digestion.

Sequenced with sAPPb F2 and F3 primers. Miniprep #15 contains both inserts in the correct orientation.

EXAMPLE 6

Construction of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695)

1. PCR of APP (664-695):

The starting material was the pRBR186 plasmid (Figure 22A).

PCR of APP (664-695)
1 ng pRBR186
50 nM 5' oligo (5'-TGCCCCGCGCGGCCGCGCGATGCTGCCCGG-3') (SEQ ID NO:59) NotI site underlined
50 nM 3' oligo (5'-(p)ATGGTGTGGCGGCCGCAGACGCCGCTGTCACC-3') (SEQ ID NO:60) NotI site underlined
1 µL Roche PCR nucleotides
5 µL 10X Expand buffer
40 µL water
1 µL Expand

Cycle: (94°C, 5 min) – 25X (94°C, 30 sec; 42°C, 1 min; 72°C, 2 min) – 72°C x 6 min – 4°C hold.

The ~100 bp fragment was gel purified (4% agarose, 1X TBE gel)

The gel-purified fragment was ligated into NotI digested, Shrimp Alkaline Phosphatase-treated pcDNA3.1 zeo (+) (Invitrogen). The presence of the insert and its orientation was confirmed by sequencing.

2. PCR of APP(1-651):
1 ng pRBR186
50 nM 5' oligo (5'-(p)AGCGCACAAGCTTCCCCGCGCAGGGTCGCGATGCTG-3') (SEQ ID NO:61) HindIII site underlined, Met(1) ATG of APP in bold
50 nM 3' oligo (5'-(p)GGATGTAAGCTTTTTCTTCTTCAGCATCACCAAGG-3') (SEQ ID NO:62) HindIII site is underlined

1 µL Roche PCR nucleotides
5 µL 10X Expand buffer
40 µL water
1 µL Expand

Cycle: (94°C, 5 min) -- 25X (94°C, 30 sec; 37°C, 1 min; 72°C 2.5 min) – 72°C x 6 min – 4°C hold

- The amplified fragment was isolated on an agarose gel. The fragment was purified from the gel using Qiaquick Gel purification columns. The fragment was digested with HindIII. The amount of the fragment was too small to subclone, so the PCR was repeated using 1 µL of the amplified fragment and carrying out 5 reactions simultaneously.

- The fragments were purified from these reactions using a QiaQuick PCR purification kit. The fragments were eluted in 30 µL and digested with HindIII for 2 hours. The digested fragments were then gel purified.

- The purified fragments were ligated to pcDNA3.1 zeo (+) APP(664-695) that had been digested with HindIII and SAP treated. This gave the intermediate plasmid pcDNA3.1 zeo (+) APP(1-651)/APP(664-695).

3. PCR of (M1L) TAT:
The starting material was NL4-3 viral plasmid (Figure 22B).

PCR reaction:
1 ng NL4-3 viral plasmid
50 nM TAT 5' RV Met-Leu PCR primer (5'-TGCAGATATCCTGGAGCCAGTAGATCCTAGAC-3') (SEQ ID NO:63)
EcoRV site underlined, Met-Leu mutation in bold
50 nM TAT 3' RV Met-Leu PCR primer
(5'-GCTGGATATCCTCTGCTTTGATAGAGAAGC-3') (SEQ ID NO:64)
EcoRV site underlined
1 µL PCR dNTPs

5 µL PCR 10X buffer with MgCl2
40 µL water
1 µL Expand polymerase

Cycle:
94°C for 5 min
[30 sec 94°C, 1 min 42°C, 1 min 72°C] x 25 cycles
5 min at 72C
hold at 4°C

- The insert was purified over QiaQuick PCR purification column
- The entire reaction was digested with 30 units EcoRV for 3 hours
- The ~200 bp insert was gel purified.
- pcDNA3.1 zeo (+) APP(1-651)/APP(664-695) was digested with EcoRV, and then SAP treated
- The Met1-Leu TAT fragment was ligated to pcDNA3.1 zeo (+) APP(1-651)/APP(664-695).

A map of the resulting plasmid is shown in Figure 22C.

EXAMPLE 7

Design of novel expression vector for expression of APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695)

PURPOSE: To provide a low level expression of APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695), a prokaryotic selectable marker that is NOT ampicillin (read-through of the β-lactamase gene is sometimes a problem), and a eukaryotic selectable marker that is NOT zeocin (zeo is the marker for the reporter plasmids in some embodiments).

METHODS

1. The dEYFP gene was removed from pd2EYFP (Clontech, Palo Alto, CA) using BamHI and NotI. The 5' overhangs was filled in using Klenow, and the plasmid was re-circularized.

pd2EYFP plasmid was digested with BamHI, NotI.
Ran reaction on 1% agarose gel. Digestion was complete. Cut out 3.4 kb band.
Purified using Qiagen Gel Extraction Kit.
Klenow fill-in:
~4 μg plasmid backbone
7.5 μL NEB buffer 2
33 μM each dNTP (diluted from Roche PCR dNTPs)
water to 75 μL
4 μL Roche Klenow fragment (4 units)

Incubated at room temperature for 15 minutes
Heat inactivated at 70°C

Took 1 μL fill-in reaction.
Diluted to 8 μL with water
Added 2 μL 5X DNA buffer
Added 10 μL 2X Ligation Buffer
Added 1 μL T4 DNA ligase

Incubated at room temperature

Transformed 2 μL ligation into Invitrogen maximum efficiency DH5alpha competent cells.

Plated out on Kanamycin plates. Lots of colonies.

2. The RSV promoter from pREP4 (Invitrogen) was excised using BglII and HindIII and cloned into the re-circularized plasmid.

Digested 5 μg pREP4 with HindIII
Purified using Qiagen PCR purification kit
Digested with BglII.

The RSV promoter fragment was gel purified and cloned into the re-circularized plasmid.

3. The resulting expression vector (pRSV Kan/Neo res; Figure 23) has the eukaryotic RSV promoter 5' to the pd2EYFP polylinker, SV40 driving neo and kanamycin prokaryotic selection, and a pUC ori for high levels of replication in bacteria.

EXAMPLE 8

Use of APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695) in HeLa cells with a β-galactosidase reporter gene

The following demonstrates the practice of the present invention with the APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695) fusion protein (SEQ ID NO:2) and β-galactosidase as a reporter gene. P4-R5 cells are HeLa cells that contain a stably integrated β-galactosidase reporter gene under the control of the HIV1 LTR.

Materials:

1.) Cells: P4-R5 cells

2.) DNA: 0.78 μg/μL pcDNA3.1 zeo (+)APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695)

3.) Transfection reagents: FUGENE®

4.) Media: OPTIMEM®
cDMEM (-)phenol red /10% FBS

5.) Compounds: Compound X (β-secretase inhibitor) 10 mM
L-875,532 (γ-secretase inhibitor) 10 mM

Protocol:

Day 1

1.) Cell count on P4-R5 cells = 7.6×10^5 cells per ml in cDMEM (-)PR.
Seeded sterile white luminometer TC plates with the following cell numbers:

	10 ml
5×10^3/well	= 0.75 ml in 9.25 ml media
1.0×10^4/well	= 1.5 ml in 8.5 ml media

Seeded 100 µL cells per well.
Incubate overnight at 37°C, 5% CO_2.

Day 2

2.) Made up media with appropriate dilutions of inhibitors.

On no-inhibitor controls, added 100 µL of cDMEM with 1% DMSO
On wells with Compound X, added 10 µM inhibitor in cDMEM
On wells with L-875,532, added 10 µM inhibitor in cDMEM

3.) Prior to transfection, pulled off media on P4-R5 cells and replaced with media -/+ inhibitor.

FUGENE® transfection:

For FUGENE® transfection:

4.) Added 600 µL of OPTIMEM® to sterile EPPENDORF® tube and carefully added 30 µL FUGENE® to media, without touching walls of tube. Incubated at room temperature for 5 minutes.

In separate EPPENDORF® tubes, added each DNA.

Added FUGENE®/OPTIMEM® dropwise to DNA; incubated at room temperature for 15 minutes.

Added 15 µL /well of DNA/FUGENE®/OPTIMEM® dropwise to media in appropriate wells on P4-R5 cells, swirling to mix.

TABLE 4

Transfection number	Conc of DNA	Vol of DNA	Vol of FUGENE®	Vol of sterile OPTIMEM®
APP-ML-Tat-APPct	0.78 µg/µL	5 µg = 6.5 µL	30 µL of FUGENE®	600 µL of OPTIMEM®
pUCd5TAT	1.24 µg/µL	5 µg = 4.0 µL	30 µL of FUGENE®	600 µL of OPTIMEM®
p243-4	0.56 µg/µL	5 µg = 9 µL	30 µL of FUGENE®	600 µL of OPTIMEM®

In Table 4, "APP-ML-Tat-APPct" refers to pcDNA3.1 zeo (+)APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695). "pUCd5TAT" is an expression vector that

serves as a positive control for TAT expression, since it is a construct in which TAT is under the control of a strong, constitutive promoter (see Figure 24). "p243-4" is a control expression vector that directs the expression of APP.

5.) Plates were transferred to an incubator and incubated for 48 hours to allow expression and processing of the proteins.

Day 4

6.) The protocol below was followed for lysis of the cells and measurement of β-galactosidase in the cell lysates.

Measurement of β-galactosidase in lysates of transfected cells.

1. Removed TROPIX® chemiluminescence kit(s) from cold room, allowed to come to room temperature in a 37°C water bath.
2. β-galactosidase standards were prepared:

Made 1:5000 dilution of β-galactosidase stock (1 mg/ml) in lysis buffer. Did 2 fold dilutions.

3. Diluted TROPIX® substrate 1:25 into buffer. (Made enough for 100 µL /well).
4. Added to reservoir and added 100 µL/well.
5. Added 10 µL of β-galactosidase standards to column 12 on plate and incubated in dark for 1 hour.
6. Read immediately in luminometer using standard file. Filled in required fields, read plate.

The results are shown in Figure 33. Figure 33 demonstrates that the present invention was able to identify both the β-secretase inhibitor (Compound X) and the γ-secretase inhibitor (L-875,532). In Figure 33, "APP-tat-ct32" refers to pcDNA3.1 zeo (+)APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695). Although not indicated in Figure 33, the results for the controls were as expected: a large transactivation of the LTR by pUCd5TAT was observed which was not affected by either inhibitor. No transactivation was seen with p243-4.

EXAMPLE 9

Comparison of the use of APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695) and APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695) with a β-galactosidase reporter gene

The following shows a side-by-side comparison of the practice of the present invention with the APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695) fusion protein (SEQ ID NO:2) and the APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695) fusion protein (SEQ ID NO:4). P4-R5 cells are HeLa cells that contain a stably integrated β-galactosidase reporter gene under the control of the HIV1 LTR.

Materials:

1.) Cells: P4-R5 cells

2.) DNA: 0.78 μg/μL pcDNA3.1 neo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695)

0.812 μg/μL pcDNA3.1 neo (+) APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695)

1.24 μg/μL pUCd5TAT

0.56 μg/μL p243-4

3.) Transfection reagents: FUGENE®

4.) Media: OPTIMEM®

cDMEM (-)phenol red /10% FBS

5.) Compounds: Compound X (β-secretase inhibitor) 10 mM

L-875,532 (γ-secretase inhibitor) 10 mM

Day 1

1.) Cell count on P4-R5 cells = 5 x 10^5 cells per ml in cDMEM (-)PR.

Seeded sterile white luminometer TC plates with the following cell numbers:

10 ml

5 x 10^3/well = 4.0 ml in 36.0 ml media

Diluted 1:1 into media and seeded one plate at 2.5 x 10^3/well.

Seeded 100 μL cells per well.

Incubated overnight at 37°C, 5% CO_2.

Day 2

2.) Made up media with appropriate dilutions of inhibitors.

On no-inhibitor controls, added 100 µL of cDMEM with 1% DMSO

On wells with Compound X, added titration curve from 100 µM inhibitor in cDMEM (-)PR.

On wells with L-875,532, added titration curve from 100 µM inhibitor in cDMEM (-)PR.

3.) Prior to transfection, pulled off media on P4-R5 cells and replaced with media -/+ inhibitor.

FUGENE® transfection:

4.) Added volume of OPTIMEM® to sterile EPPENDORF® tube and carefully added correct volume of FUGENE® to media, without touching walls of tube. Incubated at room temperature for 5 minutes.

In separate EPPENDORF® tubes, added each DNA, as outlined below.

Added FUGENE®/OPTIMEM® dropwise to DNA; incubated at room temperature for 15 minutes.

Added 15 µL/well of DNA/FUGENE®/OPTIMEM® dropwise to media in appropriate wells on P4-R5 cells, swirling to mix.

TABLE 5

Transfection number	Conc of DNA	Vol of DNA	Vol of FUGENE®	Vol of sterile OPTIMEM®
1.) APP-ML-Tat-APPct (Sw)	0.78 µg/µL	10 µg = 13 µL	60 µL of FUGENE®	1200 µL of OPTIMEM®
2.) APP-ML-Tat-APPct (WT)	0.812 µg/µL	10 µg = 12.2 µL	60 µL of FUGENE®	1200 µL of OPTIMEM®
3.) pUCd5TAT	1.24 µg/µL	5 µg = 4.0 µL	30 µL of FUGENE®	600 µL of OPTIMEM®
4.) p243-4	0.56 µg/µL	5 µg = 9 µL	30 µL of FUGENE®	600 µL of OPTIMEM®

In Table 5, "APP-ML-Tat-APPct (Sw)" refers to pcDNA3.1 neo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695). "APP-ML-Tat-APPct (WT)" refers to pcDNA3.1 neo (+) APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695). "pUCd5TAT" is an expression vector that serves as a positive control for TAT expression, since it is a construct in which TAT is under the control of a strong, constitutive promoter (see Figure 24). "p243-4" is a control expression vector that directs the expression of APP.

5.) Plates were transferred to an incubator and incubated for 36 hours to allow expression and processing of proteins.

Day 4

6.) The protocol below was followed for lysis of the cells and measurement of β-galactosidase in the cell lysates.

Measurement of β-galactosidase in lysates of transfected cells.

1. Removed TROPIX® chemiluminescence kit(s) from cold room, allowed to come to room temperature in a 37°C water bath.
2. β-galactosidase standards were prepared:
Made 1:5000 dilution of β-galactosidase stock (1 mg/ml) in lysis buffer. Did 2 fold dilutions.
3. Diluted TROPIX® substrate 1:25 into buffer. (Made enough for 120 µL /well).
4. Added to reservoir and added 120 µL/well.
5. Added 10 µL of β-galactosidase standards to column 12 on plate and incubated in dark for 1 hour.
6. Read immediately in luminometer using standard file. Filled in required fields, read plate.

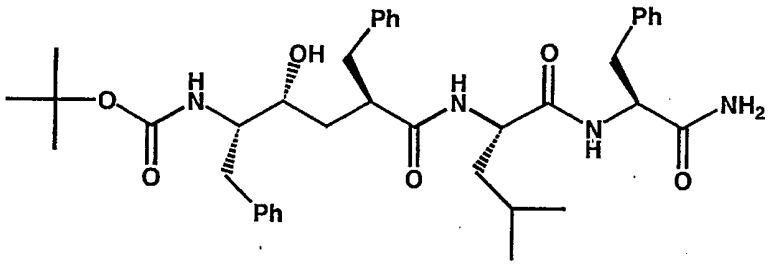
The results are shown in Figure 34. In Figure 34, "APP(NFEV)HAMycFLAG" refers to a protein that is a variant of APP in which NFEV is present at the β-secretase cleavage site and there are epitope tags in the amino terminal portion of the protein

but there is no transcription factor fused to APP. "APP(Sw)tat-ct32" refers to pcDNA3.1 neo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695). "APP(WT)tat-ct32" refers to pcDNA3.1 neo (+) APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695). Figure 34 shows that the Swedish version and the wild-type version of APP appear to work about equally well in the assay.

EXAMPLE 10

L-685,458

L-685,458 is a γ-secretase inhibitor having the following structure:



L-685,458 contains an hydroxyethylene dipeptide isostere and is thought to function as a transition state analog mimic of aspartyl proteases (Shearman et al., 2000, Biochemistry 39:8698-8704). L-685,458 was prepared as follows: {1S-Benzyl-4R-[1-(1S-carbamoyl-2-phenylethylcarbamoyl)-1S-3-methylbutylcarbamoyl]-2R-hydroxy-5-phenylpentyl}carbamic acid tert-butyl ester (L-685,458) was prepared by the coupling of 2R-benzyl-5S-tert-butoxycarbonylamino-4R-(tert-butyldimethylsilanyloxy)-6-phenylhexanoic acid (Evans et al., 1985, J. Org. Chem. 50:4615-4625) with Leu-Phe-NH2 followed by deprotection with tetrabutylammonium fluoride. The synthesis of {1S-benzyl-4R-[1-(1S-carbamoyl-2-phenylethylcarbamoyl)-1S-3-methylbutylcarbamoyl]-2S-hydroxy-5-phenylpentyl}carbamic acid tert-butyl ester (L-682,679) has been described previously (De Solms et al., 1991, J. Med. Chem. 34:2852-2857). {1S-Benzyl-4R-[1-(1S-carbamoyl-2-phenylethylcarbamoyl)-1S-3-methylbutylcarbamoyl]-2-oxo-5-phenylpentyl}carbamic acid tert-butyl ester (L-684,414) was prepared by pyridinium dichromate-mediated oxidation of L-682,679.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are intended to fall within the scope of the appended claims.

Various publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

WHAT IS CLAIMED IS:

1. A DNA molecule comprising a nucleotide sequence encoding a fusion protein comprising amino acids 589-651 selected from the group consisting of wild type APP_{695}, the Swedish version of APP_{695} and the NFEV (SEQ ID NO:40) version of APP_{695} and a transcription factor where the transcription factor is fused in frame to the carboxyl terminus of amino acids 589-651.

2. The DNA molecule of claim 1 where amino acids 589-651 contain a K612V mutation.

3. The DNA molecule of claim 1 where the nucleotide sequence further encodes amino acids 664-695 of APP_{695} wherein amino acids 664-695 are fused in frame to the carboxyl terminus of the transcription factor.

4. The DNA molecule of claim 1 where the transcription factor is selected from the group consisting of: HIV-1 TAT, Gal4-VP16, the entire Gal4 protein, LexA-VP16, E12, E47, Twist, Papillomavirus E2, EBV Zta, BIV TAT, HIV-2 TAT, or SIV TAT.

5. The DNA molecule of claim 3 where the fusion protein is selected from the group consisting of: APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:4); APP(1-651)wt, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:8); APP(1-651)wt, TATexonI(M1L) APP (664-695) (SEQ ID NO:12); and APP(1-651)wt, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:16).

6. The DNA molecule of claim 3 where the fusion protein is selected from the group consisting of: APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:2); APP(1-651)SW, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:6); APP(1-651)SW, TATexonI(M1L) APP (664-695) (SEQ ID NO:10); and APP(1-651)SW, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:14).

7. The DNA molecule of claim 3 where the fusion protein is selected from the group consisting of: APP(1-651)NFEV, K612V-TATexonI(M1L)

APP (664-695) (SEQ ID NO:23) and APP(1-651)NFEV, K612V, GAL4-VP16(M1L) APP (664-695) (SEQ ID NO:25).

8. An expression vector comprising the DNA molecule of claim 1.

9. A eukaryotic cell comprising the DNA molecule of claim 1.

10. The cell of claim 9 further comprising a reporter gene where the reporter gene is under the control of a regulatory DNA sequence that is capable of being activated by the transcription factor.

11. A method of identifying a substance that inhibits APP processing comprising:

(a) providing a recombinant eukaryotic cell which:

(i) expresses a fusion protein comprising amino acids 589-651 selected from the group consisting of wild type APP695, the Swedish version of APP695 and the NFEV (SEQ ID NO:40) version of APP695 and a transcription factor where the transcription factor is fused in frame to the carboxyl terminus of amino acids 589-651; and

(ii) comprises a reporter gene operably linked to a regulatory DNA sequence which capable of being activated by the transcription factor;

(b) measuring the level of reporter gene product in the cell in the absence of the substance;

(c) adding the compound to the cell and measuring the level of reporter gene product in the cell in the presence of the substance;

where a decrease in the level of reporter gene product in the presence as compared to the absence of the substance indicates that the substance inhibits APP processing.

12. The method of claim 11 where amino acids 589-651 contain a K612V mutation.

13. The method of claim 11 where the transcription factor is selected from the group consisting of: HIV-1 TAT, Gal4-VP16, the entire Gal4

protein, LexA-VP16, E12, E47, Twist, Papillomavirus E2, EBV Zta, BIV TAT, HIV-2 TAT, or SIV TAT.

14. The method of claim 11 where the fusion protein is selected from the group consisting of: APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:4); APP(1-651)wt, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:8); APP(1-651)wt, TATexonI(M1L) APP (664-695) (SEQ ID NO:12); and APP(1-651)wt, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:16).

15. The method of claim 11 where the fusion protein is selected from the group consisting of: APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:2); APP(1-651)SW, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:6); APP(1-651)SW, TATexonI(M1L) APP (664-695) (SEQ ID NO:10); and APP(1-651)SW, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:14).

16. The method of claim 11 where the fusion protein is selected from the group consisting of: APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:23) and APP(1-651)NFEV, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:25).

17. A method of identifying a substance that inhibits APP processing comprising:

(a) providing a recombinant eukaryotic cell which:

(i) expresses a fusion protein comprising an amino acid sequence from APP that is capable of being cleaved by both β-secretase and γ-secretase and a transcription factor where the transcription factor is fused in frame to the amino acid sequence from APP; and

(ii) comprises a reporter gene operably linked to a regulatory DNA sequence which capable of being activated by the transcription factor;

(b) measuring the level of reporter gene product in the cell in the absence of the substance;

(c) adding the compound to the cell and measuring the level of reporter gene product in the cell in the presence of the substance;

where a decrease in the level of reporter gene product in the presence as compared to the absence of the substance indicates that the substance inhibits APP processing.

18. The method of claim 17 where the amino acid sequence from APP comprises an amino acid sequence selected from the group consisting of 589-651 of APP_{695}, 589-651 of the Swedish version of APP_{695}, and 589-651 of the NFEV version of APP_{695}.

19. The method of claim 17 where the amino acid sequence from APP contains the amino acid sequence NLDA (SEQ ID NO:38) at the β-secretase cleavage site instead of the wild-type sequence KMDA (SEQ ID NO:34).

20. The method of claim 17 where the amino acid sequence from APP contains the amino acid sequence NFEV (SEQ ID NO:40) at the β-secretase cleavage site instead of the wild-type sequence KMDA (SEQ ID NO:34).

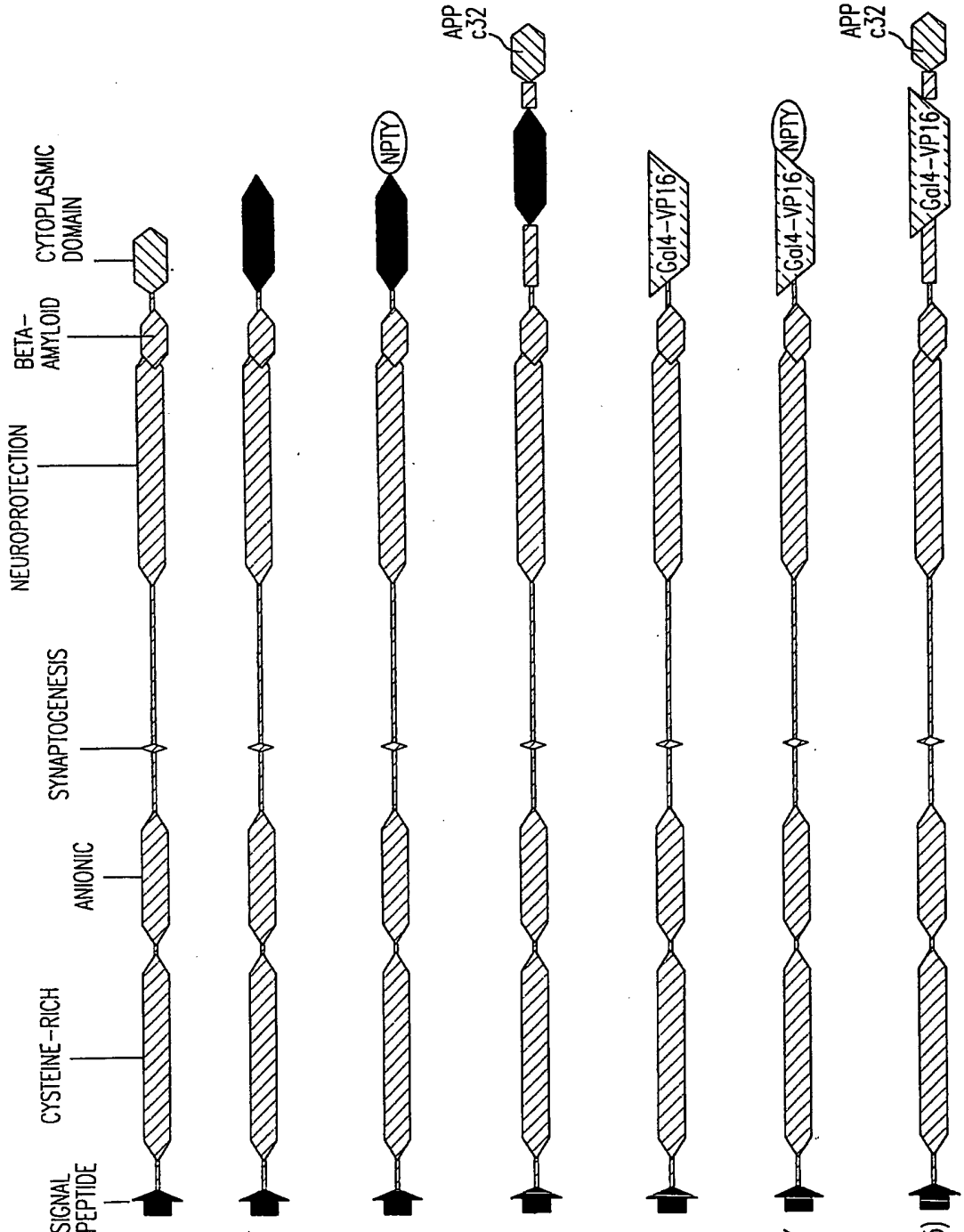


FIG. 1A

FIG. 1B

FIG. 1C

FIG. 1D

FIG. 1E

FIG. 1F

FIG. 1G

DNA sequence of APP(1-651)SW, K612V-TATexon1(M1 L) APP (664-695)

(SEQ ID NO: 1)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC
1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
```

FIG.2A

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC

1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC

1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC

1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG

1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG

1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC

1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA

1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAATCTAGA TGCAGAATTC

1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTTGC

1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG

1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG

1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC

2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC

2051 CAGGAAGTCA GCCTAAAACT GCTTGTACCA ATTGCTATTG TAAAAAGTGT

2101 TGCTTTCATT GCCAAGTTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA

2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA

2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC

2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA

2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT

2351 AG

FIG.2B

(SEQ ID NO: 2)

Amino acid sequence of APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefsldd1qpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

nldaefrhdsgyevhhqvlvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

6

~~adi~~lepvdprlepwkhpgsqpktactncyckkccfhcqvcfmtkalgisygrkkrrqrrrahqnsqthqaslskq

7 8

~~risstvaaa~~daavtpeerhlskmqqnqyenptykffeqmqn

FIG.3

DNA sequence of APP(1-651)wt, K612V-TATexon1(M1L) APP (664-695)
(SEQ ID NO: 3)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC
1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
```

FIG.4A

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAAGATGGA TGCAGAATTC
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC
2051 CAGGAAGTCA GCCTAAAACT GCTTGTACCA ATTGCTATTG TAAAAAGTGT
2101 TGCTTTCATT GCCAAGTTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG.4B

(SEQ ID NO: 4)

Amino acid sequence of APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefsldd1qpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

kmdaefrhdsgyevhhqvlvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

6

~~adi~~lepvdprlepwkhpgsqpktactncyckkccfhcqvcfmtkalgisygrkkrrqrrrahqnsqthqaslskq

7 8

~~risstvaaa~~daavtpeerhlskmqqngyenptykffeqmqn

FIG.5

DNA sequence of APP(1-651)SW, K612V-GAL4VP16(M1 L) APP (664-695)

(SEQ ID NO: 5)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC
```

FIG.6A

1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAATCTAGA TGCAGAATTC
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC
2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTTGCC
2051 GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT
2101 CTGAAGAACA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC
2151 GCTGACTAGG GCACATCTGA CAGAAGTGGA ATCAAGGCTA GAAAGACTGG
2201 AACAGCTATT TCTACTGATT TTTCCTCGAG AAGACCTTGA CATGATTTTG
2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT
2301 ACAAGATAAT GTGAATAAAG ATGCCGTCAC AGATAGATTG GCTTCAGTGG
2351 AGACTGATAT GCCTCTAACA TTGAGACAGC ATAGAATAAG TGCGACATCA
2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG
2451 AATTCCCGGG GATCTGGCCC CCCCGACCGA TGTCAGCCTG GGGGACGAGC
2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC
2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.6B

2601 TACCCCCCAC GACTCCGCCC CCTACGGCGC TCTGGATATG GCCGACTTCG

2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG

2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG

2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT

2801 TCTTTGAGCA GATGCAGAAC TAG

FIG.6C

(SEQ ID NO: 6)

Amino acid sequence of APP(1-651)SW, K612V, GAL4-VP16(delM1) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslslllynvpavaeeiqdevdellqkeqnysddvlanmisepriSygndal

mpsltetkttvellpvngefsldd1qpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

nldaefrhdsgyevhhqvlvffaedvgsnkgaiiglmvggvviatvivitlvmlkkkk~~klgtelgstspvwwns~~

~~adi~~kllssieqacdicrlkklkcskekpkcakclknnwecryspktkrspltrahltevesrlerleqlfllifpredld

6

milkmdslqdikalltglfvqdnvnkdavtdrlasvetdmpltlrqhrisatssseessnkqqrqltvsqipqdlapp

tdvslgdelhldqedvamahadaldd fdldmlqdqdspqpqftphdsapyqaldmadfefeqmftdalqidey

7 8

~~ggdiqhsgaaa~~daavtpeerhlskmqqnqyenptykffeqmqn

FIG.7

DNA sequence of APP(1-651)wt, K612V, GAL4-VP16(deIM1) APP (664-695)

(SEQ ID NO: 7)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
```

FIG.8A

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC

1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC

1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC

1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC

1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG

1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG

1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC

1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA

1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAAGATGGA TGCAGAATTC

1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTTGC

1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG

1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG

1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC

2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTTGCC

2051 GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT

2101 CTGAAGAACA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC

2151 GCTGACTAGG GCACATCTGA CAGAAGTGGA ATCAAGGCTA GAAAGACTGG

2201 AACAGCTATT TCTACTGATT TTTCCTCGAG AAGACCTTGA CATGATTTTG

2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT

2301 ACAAGATAAT GTGAATAAAG ATGCCGTCAC AGATAGATTG GCTTCAGTGG

2351 AGACTGATAT GCCTCTAACA TTGAGACAGC ATAGAATAAG TGCGACATCA

2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG

2451 AATTCCCGGG GATCTGGCCC CCCCGACCGA TGTCAGCCTG GGGGACGAGC

2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC

2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.8B

2601 TACCCCCCAC GACTCCGCCC CCTACGGCGC TCTGGATATG GCCGACTTCG

2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG

2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG

2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT

2801 TCTTTGAGCA GATGCAGAAC TAG

FIG.8C

(SEQ ID NO: 8)

Amino acid sequence of APP(1-651)wt, K612V, GAL4-VP16(delM1) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefsldd1qpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

kmdaefrhdsgyevhhqklvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

~~adi~~kllssieqacdicrlkklkcskekpkcakclknnwecryspktkrspltrahltevesrlerleqlfllifpredld

6

milkmdslqdikalltqlfvqdnvnkdavtdrlasvetdmpltlrqhrisatssseessnkqqrqltvsqipqdlapp

tdvslgdelhldgedvamahadaldd fdldmlgdgdspgpgftphdsapygaldmadfefeqmftdalgidey

7 8

~~ggdiqhsgaaa~~daavtpeerhlskmqqnqyenptykffeqmqn

FIG.9

DNA sequence of APP(1-651)SW, TATexon1(M1 L) APP (664-695)

(SEQ ID NO: 9)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
```

FIG.10A

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC

1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC

1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC

1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC

1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG

1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG

1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC

1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA

1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAATCTAGA TGCAGAATTC

1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTTGC

1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG

1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG

1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC

2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC

2051 CAGGAAGTCA GCCTAAAACT GCTTGTACCA ATTGCTATTG TAAAAAGTGT

2101 TGCTTTCATT GCCAAGTTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA

2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA

2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC

2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA

2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT

2351 AG

FIG.10B

(SEQ ID NO: 10)

Amino acid sequence of APP(1-651)SW, TATexonI(M1L) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslslllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefsldd1qpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

nldaefrhdsgyevhhqklvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

6

~~adi~~lepvdprlepwkhpgsqpktactncyckkccfhcqvcfmtkalgisygrkkrrqrrrahqnsqthqaslskq

7 8

~~risstvaaa~~daavtpeerhlskmqqngyenptykffeqmqn

FIG.11

DNA sequence of APP(1-651)wt, TATexon1(M1L)-APP (664-695)

(SEQ ID NO: 11)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
```

FIG.12A

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC

1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC

1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC

1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC

1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG

1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG

1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC

1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA

1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAAGATGGA TGCAGAATTC

1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTTGC

1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG

1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG

1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC

2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC

2051 CAGGAAGTCA GCCTAAAACT GCTTGTACCA ATTGCTATTG TAAAAAGTGT

2101 TGCTTTCATT GCCAAGTTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA

2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA

2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC

2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA

2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT

2351 AG

FIG.12B

(SEQ ID NO: 12)

Amino acid sequence of APP(1-651)wt, TATexonI(M1L) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefsldd1qpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

kmdaefrhdsgyevhhqklvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

6

~~adi~~lepvdprlepwkhpgsqpktactncyckkccfhcqvcfmtkalqisygrkkrrqrrrahqnsqthqaslskq

7 8

~~risstvaaa~~daavtpeerhlskmqqngyenptykffeqmqn

FIG.13

DNA sequence of APP(1-651)SW, GAL4VP16(delMet) APP (664-695)

(SEQ ID NO: 13)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
```

FIG.14A

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC

1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC

1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC

1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC

1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG

1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG

1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC

1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA

1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAATCTAGA TGCAGAATTC

1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTTGC

1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG

1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG

1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC

2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTTGCC

2051 GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT

2101 CTGAAGAACA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC

2151 GCTGACTAGG GCACATCTGA CAGAAGTGGA ATCAAGGCTA GAAAGACTGG

2201 AACAGCTATT TCTACTGATT TTTCCTCGAG AAGACCTTGA CATGATTTTG

2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT

2301 ACAAGATAAT GTGAATAAAG ATGCCGTCAC AGATAGATTG GCTTCAGTGG

2351 AGACTGATAT GCCTCTAACA TTGAGACAGC ATAGAATAAG TGCGACATCA

2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG

2451 AATTCCCGGG GATCTGGCCC CCCCGACCGA TGTCAGCCTG GGGGACGAGC

2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC

2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.14B

2601 TACCCCCCAC GACTCCGCCC CCTACGGCGC TCTGGATATG GCCGACTTCG

2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG

2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG

2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT

2801 TCTTTGAGCA GATGCAGAAC TAG

FIG.14C

(SEQ ID NO: 14)

Amino acid sequence of APP(1-651)SW, K612V, GAL4-VP16(delM1) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefslddlqpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

nldaefrhdsgyevhhqklvffaedvgsnkgaiiglmvggvviatvivitlvmlkkkklgtelgstspvwwns

adikllssieqacdicrlkklkcskekpkcakclknnwecryspktkrspltrahltevesrlerleqlfllifpredld

6

milkmdslqdikalltqlfvqdnvnkdavtdrlasvetdmpltlrqhrisatssseessnkgqrqltvsqipqdlapp

tdvslgdelhldgedvamahadalddfdldmlgdgdspgpqftphdsapygaldmadfefeqmftdalgidey

7 8

ggdiqhsgaaadaavtpeerhlskmqqnqyenptykffeqmqn

FIG.15

DNA sequence of APP(1-651)wt, GAL4VP16(delMet) APP (664-695)

(SEQ ID NO: 15)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
```

FIG.16A

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC

1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC

1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC

1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC

1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG

1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG

1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC

1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA

1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAAGATGGA TGCAGAATTC

1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTTGC

1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG

1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG

1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC

2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTTGCC

2051 GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT

2101 CTGAAGAACA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC

2151 GCTGACTAGG GCACATCTGA CAGAAGTGGA ATCAAGGCTA GAAAGACTGG

2201 AACAGCTATT TCTACTGATT TTTCCTCGAG AAGACCTTGA CATGATTTTG

2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT

2301 ACAAGATAAT GTGAATAAAG ATGCCGTCAC AGATAGATTG GCTTCAGTGG

2351 AGACTGATAT GCCTCTAACA TTGAGACAGC ATAGAATAAG TGCGACATCA

2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG

2451 AATTCCCGGG GATCTGGCCC CCCCGACCGA TGTCAGCCTG GGGGACGAGC

2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC

2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.16B

2601 TACCCCCCAC GACTCCGCCC CCTACGGCGC TCTGGATATG GCCGACTTCG

2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG

2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG

2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT

2801 TCTTTGAGCA GATGCAGAAC TAG

FIG.16C

(SEQ ID NO: 16)

Amino acid sequence of APP(1-651)wt, K612V, GAL4-VP16(delM1) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefsldd1qpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

kmdaefrhdsgyevhhqklvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

~~adi~~kllssieqacdicrlkklkcskekpkcakclknnwecryspktkrspltrahltevesrlerleqlfllifpredld

6

milkmdslqdikalltglfvqdnvnkdavtdrlasvetdmpltlrqhrisatssseessnkqqrqltvsqipqdlapp

tdvslqdelhldqedvamahadaldd fdldmlgdqdspqpqftphdsapyqaldmadfefeqmftdalgidey

7 8

~~ggdiqhsgaaa~~daavtpeerhlskmqqnqyenptykffeqmqn

FIG.17

(SEQ ID NO: 17)

```
   1 agtttcctcg gcagcggtag gcgagagcac gcggaggagc gtgcgcgggg gccccgggag
  61 acggcggcgg tggcggcgcg ggcagagcaa ggacgcggcg gatcccactc gcacagcagc
 121 gcactcggtg ccccgcgcag ggtcgcgatg ctgcccggtt tggcactgct cctgctggcc
 181 gcctggacgg ctcgggcgct ggaggtaccc actgatggta atgctggcct gctggctgaa
 241 ccccagattg ccatgttctg tggcagactg aacatgcaca tgaatgtcca gaatgggaag
 301 tgggattcag atccatcagg gaccaaaacc tgcattgata ccaaggaagg catcctgcag
 361 tattgccaag aagtctaccc tgaactgcag atcaccaatg tggtagaagc caaccaacca
 421 gtgaccatcc agaactggtg caagcggggc cgcaagcagt gcaagaccca tccccacttt
 481 gtgattccct accgctgctt agttggtgag tttgtaagtg atgcccttct cgttcctgac
 541 aagtgcaaat tcttacacca ggagaggatg gatgtttgcg aaactcatct tcactggcac
 601 accgtcgcca aagagacatg cagtgagaag agtaccaact tgcatgacta cggcatgttg
 661 ctgccctgcg gaattgacaa gttccgaggg gtagagtttg tgtgttgccc actggctgaa
 721 gaaagtgaca atgtggattc tgctgatgcg gaggaggatg actcggatgt ctggtggggc
 781 ggagcagaca cagactatgc agatgggagt gaagacaaag tagtagaagt agcagaggag
 841 gaagaagtgg ctgaggtgga agaagaagaa gccgatgatg acgaggacga tgaggatggt
 901 gatgaggtag aggaagaggc tgaggaaccc tacgaagaag ccacagagag aaccaccagc
 961 attgccacca ccaccaccac caccacagag tctgtggaag aggtggttcg agttcctaca
1021 acagcagcca gtacccctga tgccgttgac aagtatctcg agacacctgg ggatgagaat
1081 gaacatgccc atttccagaa agccaaagag aggcttgagg ccaagcaccg agagagaatg
1141 tcccaggtca tgagagaatg ggaagaggca gaacgtcaag caaagaactt gcciaaagct
1201 gataagaagg cagttatcca gcatttccag gagaaagtgg aatctttgga acaggaagca
1261 gccaacgaga gacagcagct ggtggagaca cacatggcca gagtggaagc catgctcaat
1321 gaccgccgcc gcctggccct ggagaactac atcaccgctc tgcaggctgt tcctcctcgg
1381 cctcgtcacg tgttcaatat gctaaagaag tatgtccgcg cagaacagaa ggacagacag
1441 cacaccctaa agcatttcga gcatgtgcgc atggtggatc ccaagaaagc cgctcagatc
1501 cggtcccagg ttatgacaca cctccgtgtg atttatgagc gcatgaatca gtctctctcc
1561 ctgctctaca acgtgcctgc agtggccgag gagattcagg atgaagttga tgagctgctt
1621 cagaaagagc aaaactattc agatgacgtc ttggccaaca tgattagtga accaaggatc
1681 agttacggaa acgatgctct catgccatct ttgaccgaaa cgaaaaccac cgtggagctc
1741 cttcccgtga atggagagtt cagcctggac gatctccagc cgtggcattc ttttggggct
1801 gactctgtgc cagccaacac agaaaacgaa gttgagcctg ttgatgcccg ccctgctgcc
1861 gaccgaggac tgaccactcg accaggttct gggttgacaa atatcaagac ggaggagatc
1921 tctgaagtga agatggatgc agaattccga catgactcag gatatgaagt tcatcatcaa
1981 aaattggtgt tctttgcaga agatgtgggt tcaaacaaag gtgcaatcat tggactcatg
2041 gtgggcggtg ttgtcatagc gacagtgatc gtcatcacct tggtgatgct gaagaagaaa
2101 cagtacacat ccattcatca tggtgtggtg gaggttgacg ccgctgtcac cccagaggag
2161 cgccacctgt ccaagatgca gcagaacggc tacgaaaatc caacctacaa gttctttgag
2221 cagatgcaga actagacccc cgccacagca gcctctgaag ttggacagca aaaccattgc
2281 ttcactaccc atcggtgtcc atttatagaa taatgtggga agaaacaaac ccgttttatg
2341 atttactcat tatcgccttt tgacagctgt gctgtaacac aagtagatgc ctgaacttga
2401 attaatccac acatcagtaa tgtattctat ctctctttac attttggtct ctatactaca
2461 ttattaatgg gttttgtgta ctgtaaagaa tttagctgta tcaaactagt gcatgaatag
2521 attctctcct gattatttat cacatagccc cttagccagt tgtatattat tcttgtggtt
2581 tgtgacccaa ttaagtccta ctttacatat gctttaagaa tcgatggggg atgcttcatg
2641 tgaacgtggg agttcagctg cttctcttgc ctaagtattc ctttcctgat cactatgcat
2701 tttaaagtta aacattttta agtatttcag atgctttaga gagatttttt ttccatgact
2761 gcattttact gtacagattg ctgcttctgc tatatttgtg atataggaat taagaggata
```

FIG.18A

```
2821 cacacgtttg tttcttcgtg cctgttttat gtgcacacat taggcattga gacttcaagc
2881 ttttcttttt ttgtccacgt atctttgggt ctttgataaa gaaaagaatc cctgttcatt
2941 gtaagcactt ttacggggcg ggtggggagg ggtgctctgc tggtcttcaa ttaccaagaa
3001 ttctccaaaa caattttctg caggatgatt gtacagaatc attgcttatg acatgatcgc
3061 tttctacact gtattacata aataaattaa ataaaataac cccgggcaag acttttcttt
3121 gaaggatgac tacagacatt aaataatcga agtaattttg ggtggggaga agaggcagat
3181 tcaattttct ttaaccagtc tgaagtttca tttatgatac aaaagaagat gaaaatggaa
3241 gtggcaatat aaggggatga ggaaggcatg cctggacaaa cccttctttt aagatgtgtc
3301 ttcaatttgt ataaaatggt gttttcatgt aaataaatac attcttggag gagc
```

FIG. 18B

(SEQ ID NO: 18)

MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMH
MNVQNGKWDSDPSGTKTCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGR
KQCKTHPHFVIPYRCLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSE
KSTNLHDYGMLLPCGIDKFRGVEFVCCPLAEESDNVDSADAEEDDSDVWWGGADTDYA
DGSEDKVVEVAEEEEVAEVEEEEADDDEDDEDGDEVEEEAEEPYEEATERTTSIATTT
TTTTESVEEVVRVPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQV
MREWEEAERQAKNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLND
RRRLALENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQ
IRSQVMTHLRVIYERMNQSLSLLYNVPAVAEEIQDEVDELLQKEQNYSDDVLANMISE
PRISYGNDALMPSLTETKTTVELLPVNGEFSLDDLQPWHSFGADSVPANTENEVEPVD
ARPAADRGLTTRPGSGLTNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNK
GAIIGLMVGGVVIATVIVITLVMLKKKQYTSIHHGVVEVDAAVTPEERHLSKMQQNGY
ENPTYKFFEQMQN

FIG.18C

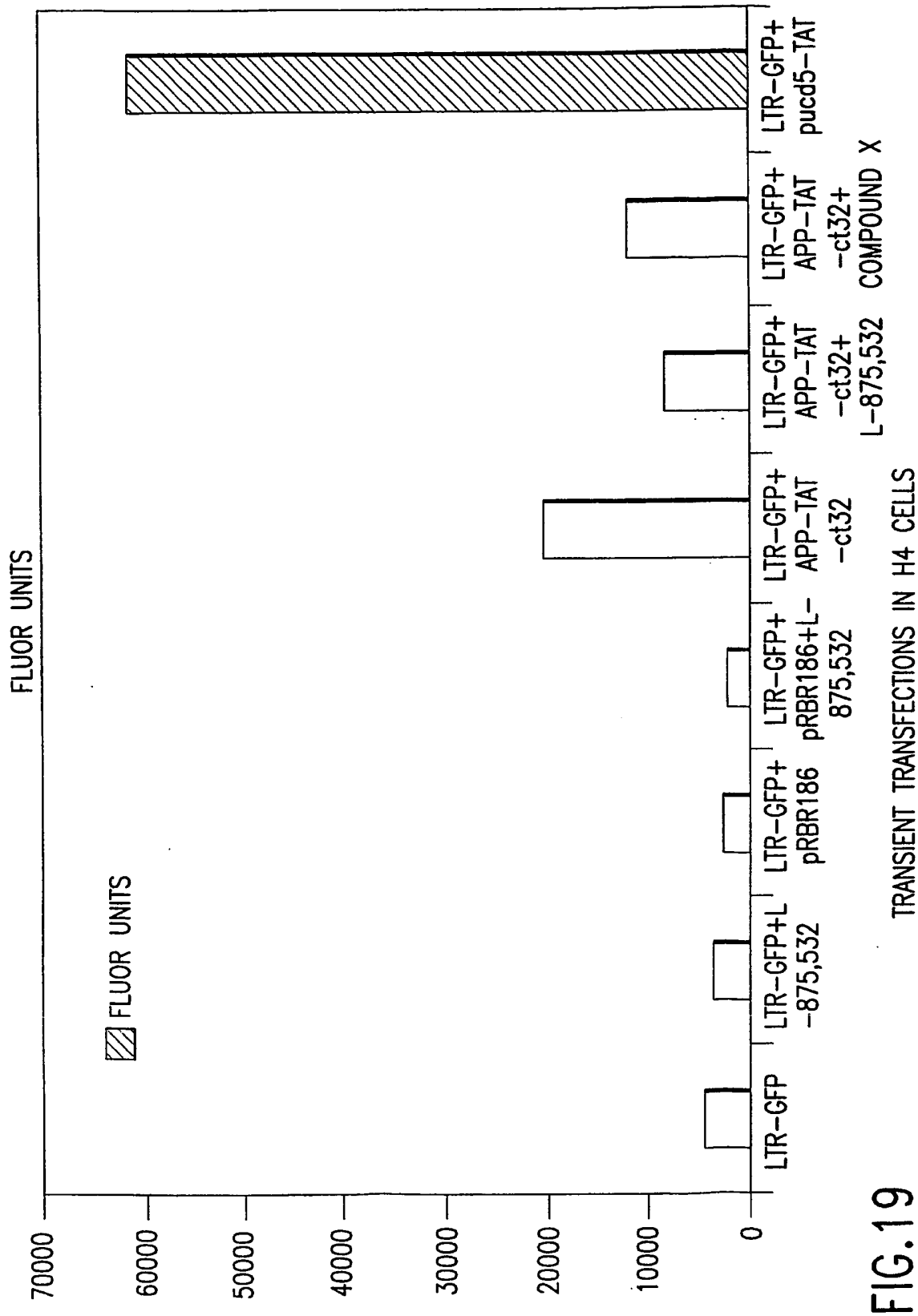

FLUOR UNITS
FLUOR UNITS
70000
60000
50000
40000
30000
20000
10000
0
LTR-GFP
LTR-GFP+L -875,532
LTR-GFP+ pRBR186
LTR-GFP+ pRBR186+L- 875,532
LTR-GFP+ APP-TAT -ct32
LTR-GFP+ APP-TAT -ct32+ L-875,532
LTR-GFP+ APP-TAT -ct32+ COMPOUND X
LTR-GFP+ pucd5-TAT
TRANSIENT TRANSFECTIONS IN H4 CELLS

FIG. 19

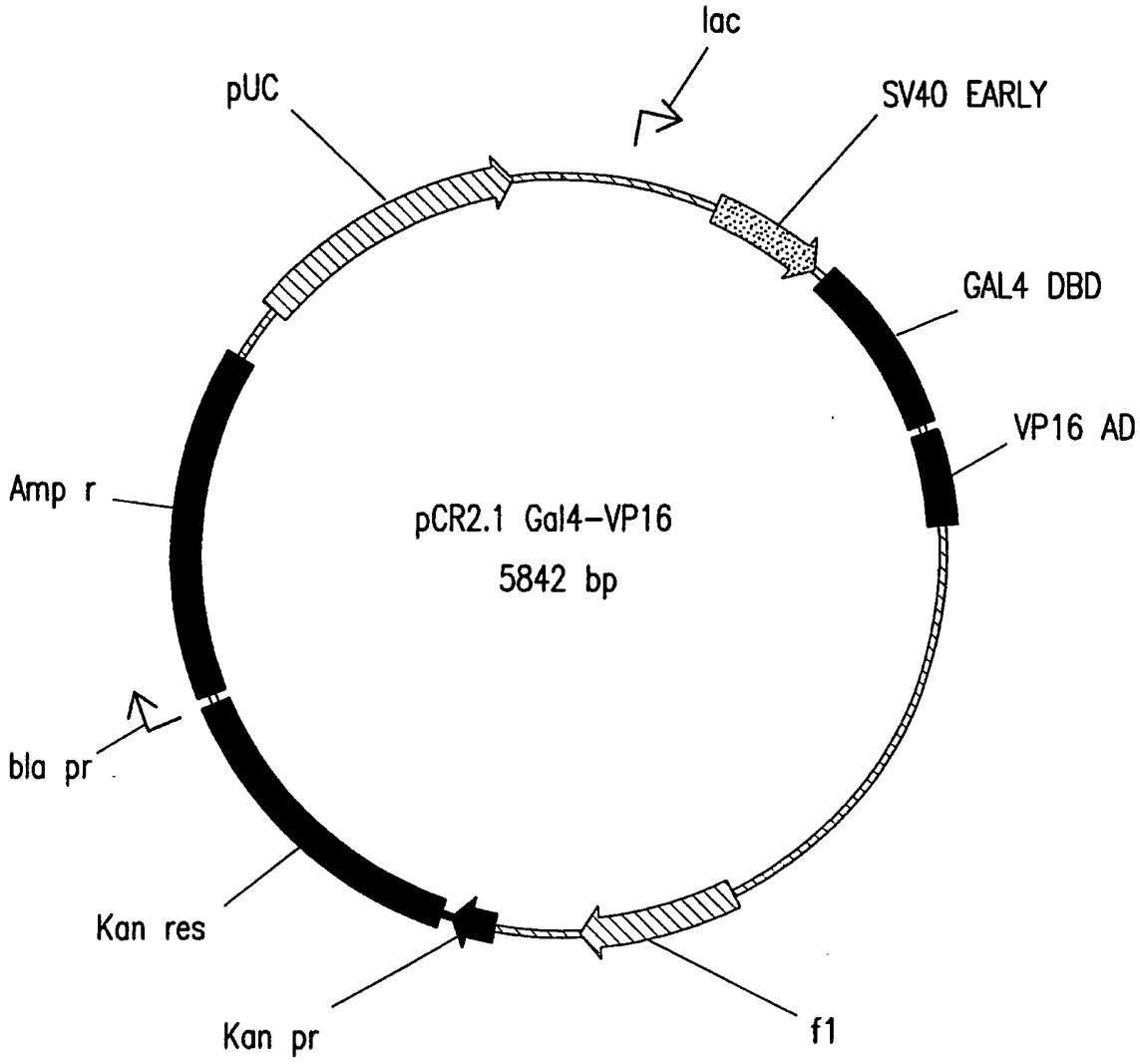

lac
pUC
SV40 EARLY
GAL4 DBD
VP16 AD
Amp r
pCR2.1 Gal4-VP16
5842 bp
bla pr
Kan res
Kan pr
f1

FIG.20

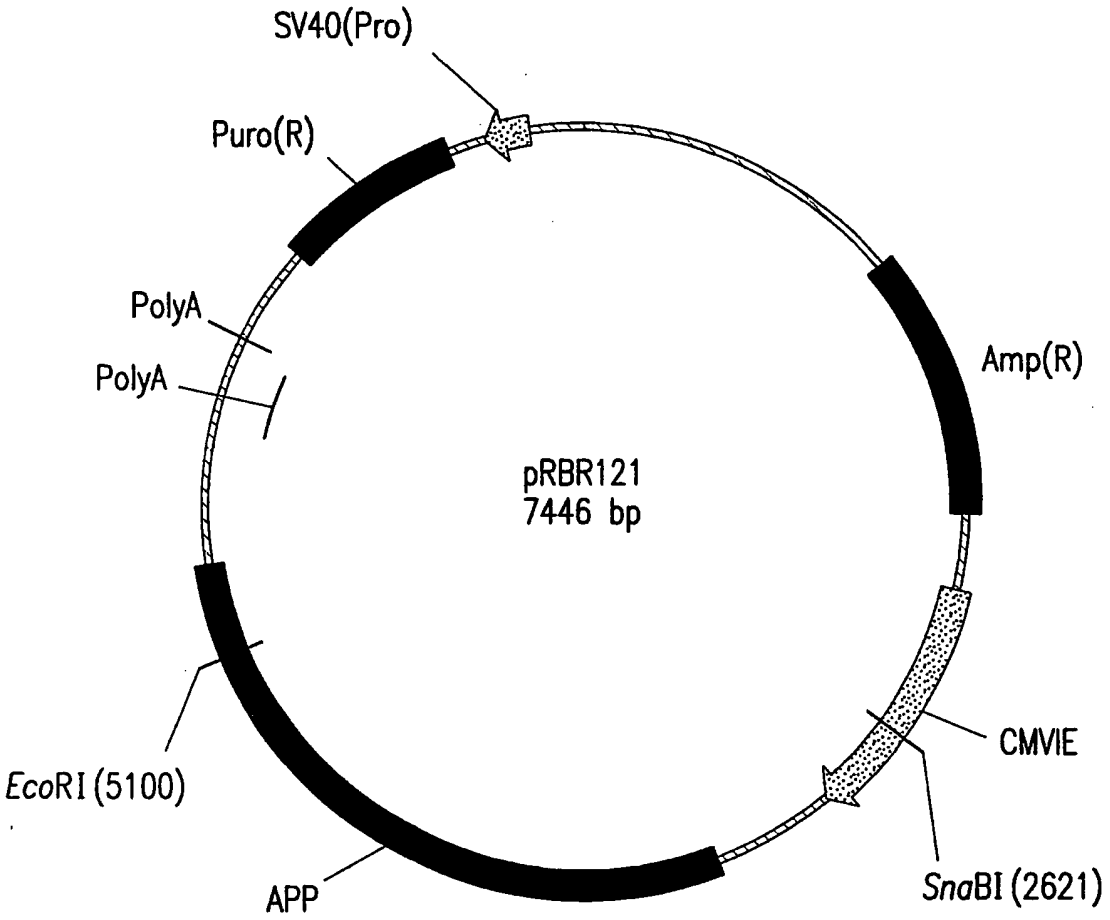

SV40(Pro)
Puro(R)
PolyA
PolyA
Amp(R)
pRBR121
7446 bp
CMVIE
EcoRI (5100)
APP
SnaBI (2621)

FIG.21A

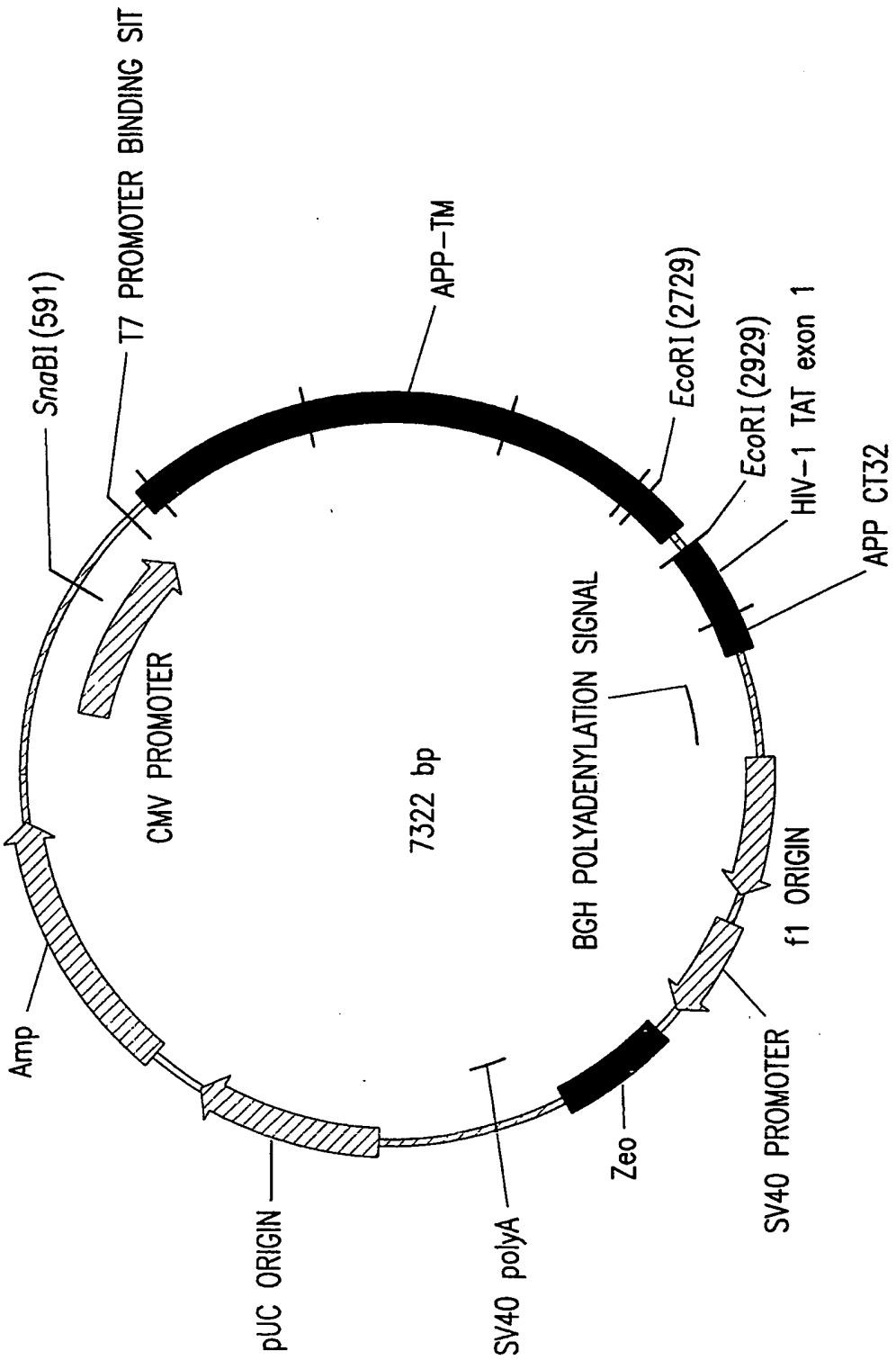

T7 PROMOTER BINDING SIT
SnaBI (591)
APP-TM
EcoRI (2729)
EcoRI (2929)
HIV-1 TAT exon 1
APP CT32
CMV PROMOTER
7322 bp
BGH POLYADENYLATION SIGNAL
f1 ORIGIN
Amp
SV40 PROMOTER
Zeo
pUC ORIGIN
SV40 polyA

FIG.21B

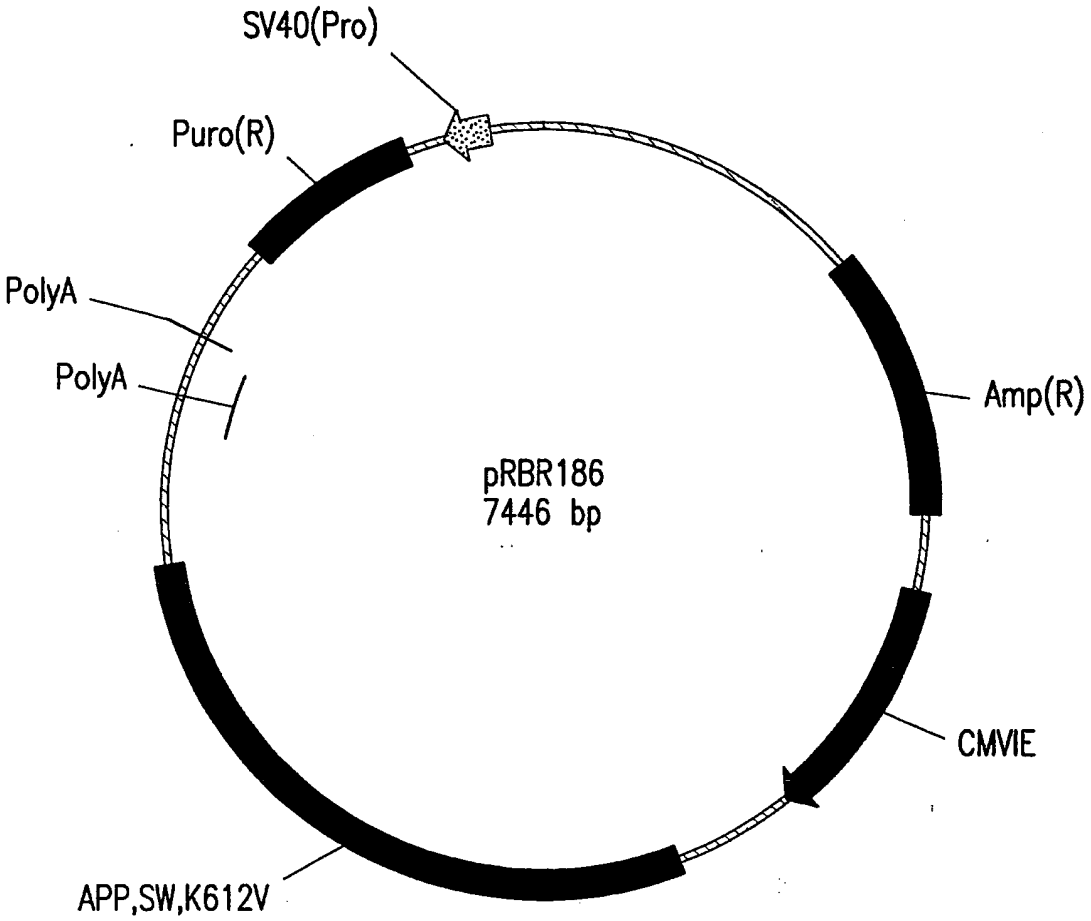

SV40(Pro)
Puro(R)
PolyA
PolyA
Amp(R)
pRBR186
7446 bp
CMVIE
APP,SW,K612V

FIG.22A

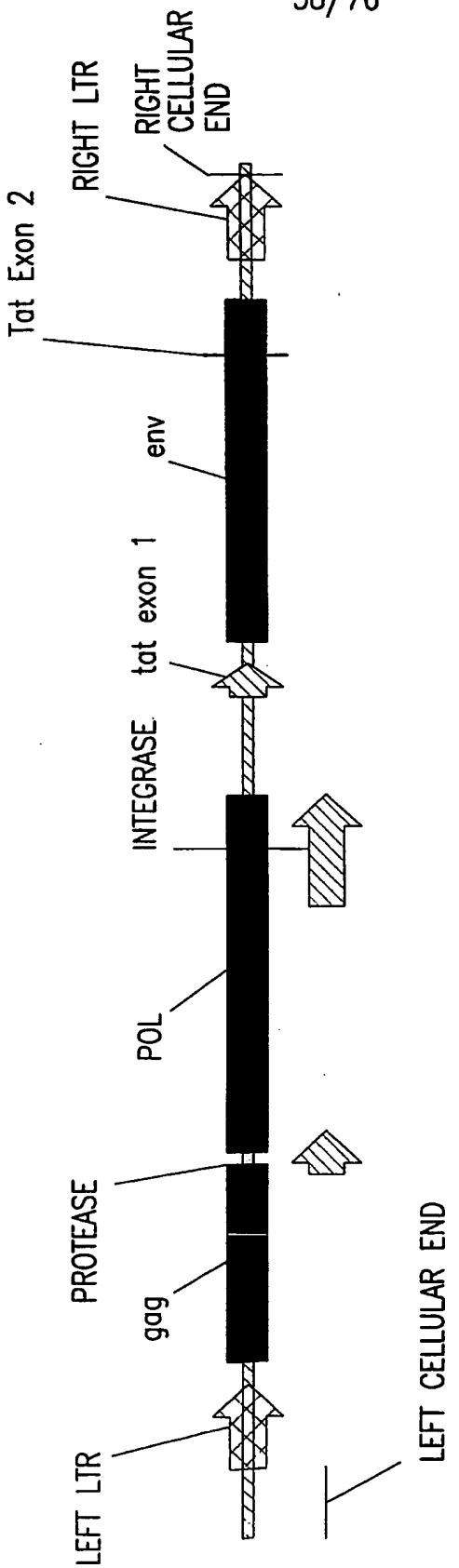

RIGHT LTR
RIGHT CELLULAR END
Tat Exon 2
env
tat exon 1
INTEGRASE
POL
PROTEASE
gag
LEFT LTR
LEFT CELLULAR END
10317 bp

FIG.22B

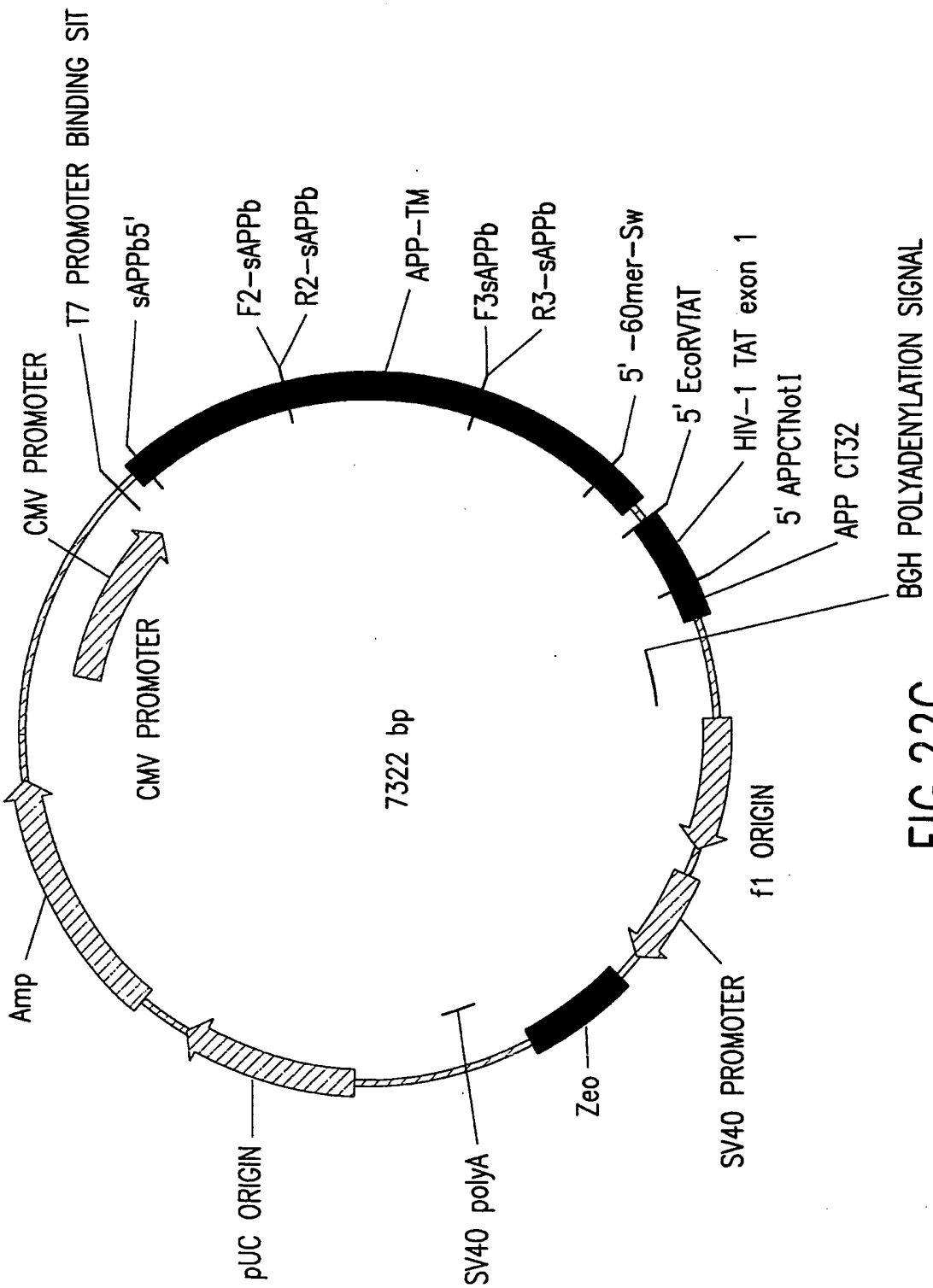

CMV PROMOTER
T7 PROMOTER BINDING SIT
sAPPb5'
F2-sAPPb
R2-sAPPb
APP-TM
F3sAPPb
R3-sAPPb
5' -60mer-Sw
5' EcoRVTAT
HIV-1 TAT exon 1
5' APPCTNotI
APP CT32
BGH POLYADENYLATION SIGNAL
CMV PROMOTER
7322 bp
f1 ORIGIN
SV40 PROMOTER
Zeo
SV40 polyA
pUC ORIGIN
Amp

FIG.22C

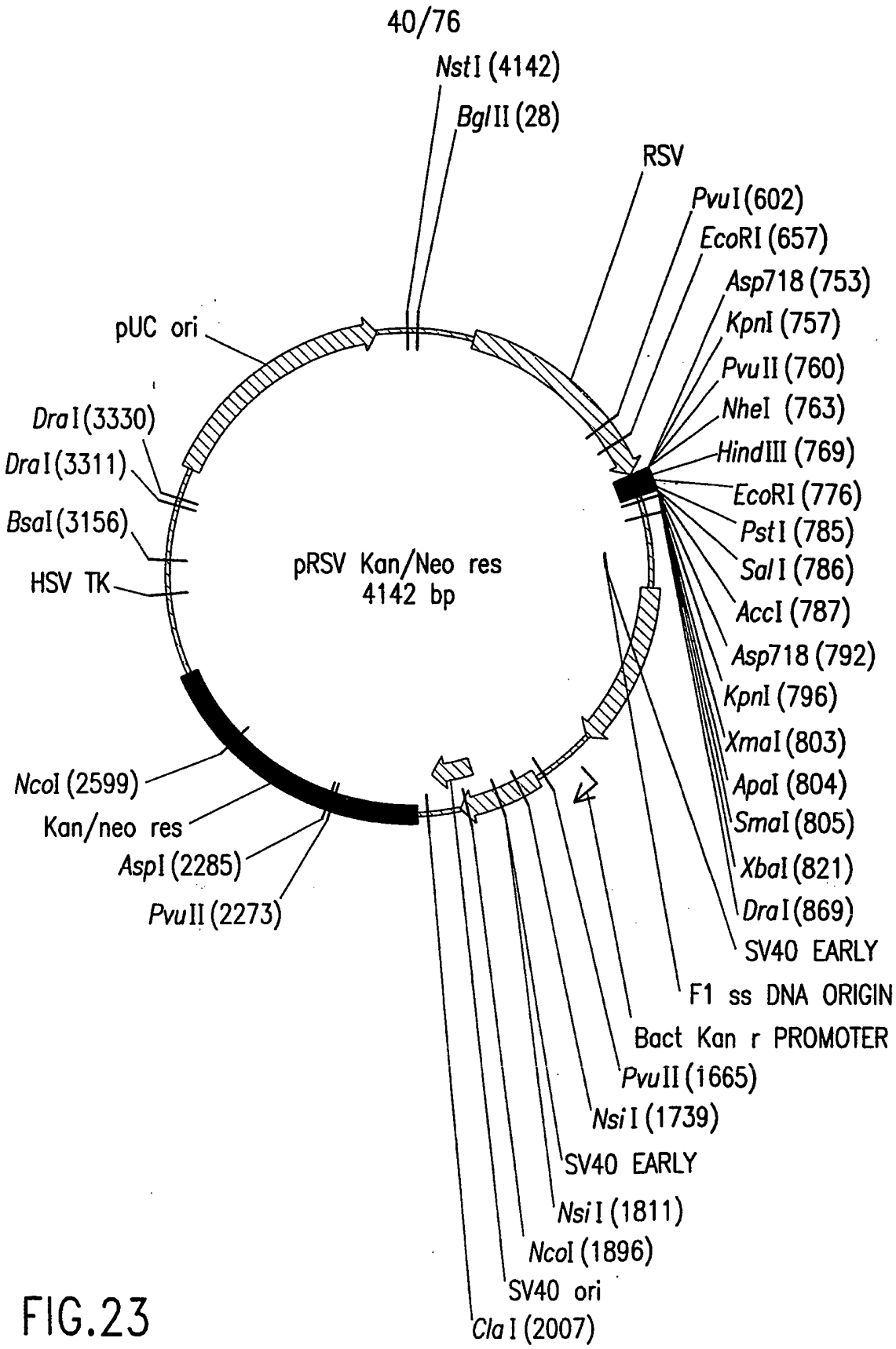
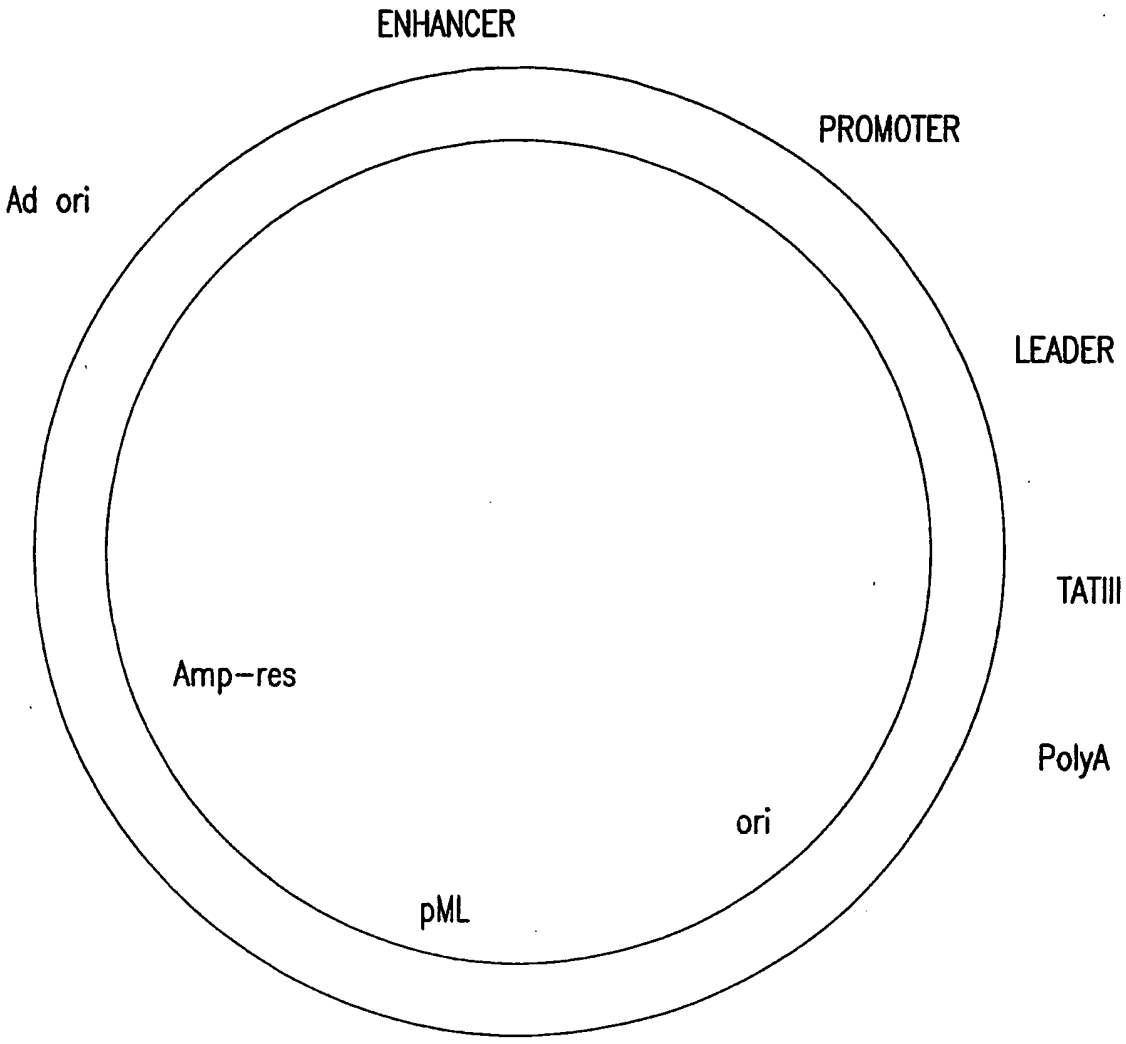

NstI (4142)
BglII (28)
RSV
PvuI (602)
EcoRI (657)
Asp718 (753)
KpnI (757)
PvuII (760)
NheI (763)
HindIII (769)
EcoRI (776)
PstI (785)
SalI (786)
AccI (787)
Asp718 (792)
KpnI (796)
XmaI (803)
ApaI (804)
SmaI (805)
XbaI (821)
DraI (869)
SV40 EARLY
F1 ss DNA ORIGIN
Bact Kan r PROMOTER
PvuII (1665)
NsiI (1739)
SV40 EARLY
NsiI (1811)
NcoI (1896)
SV40 ori
ClaI (2007)
pUC ori
DraI (3330)
DraI (3311)
BsaI (3156)
HSV TK
pRSV Kan/Neo res
4142 bp
NcoI (2599)
Kan/neo res
AspI (2285)
PvuII (2273)

FIG.23



pUCd5TAT

FIG.24

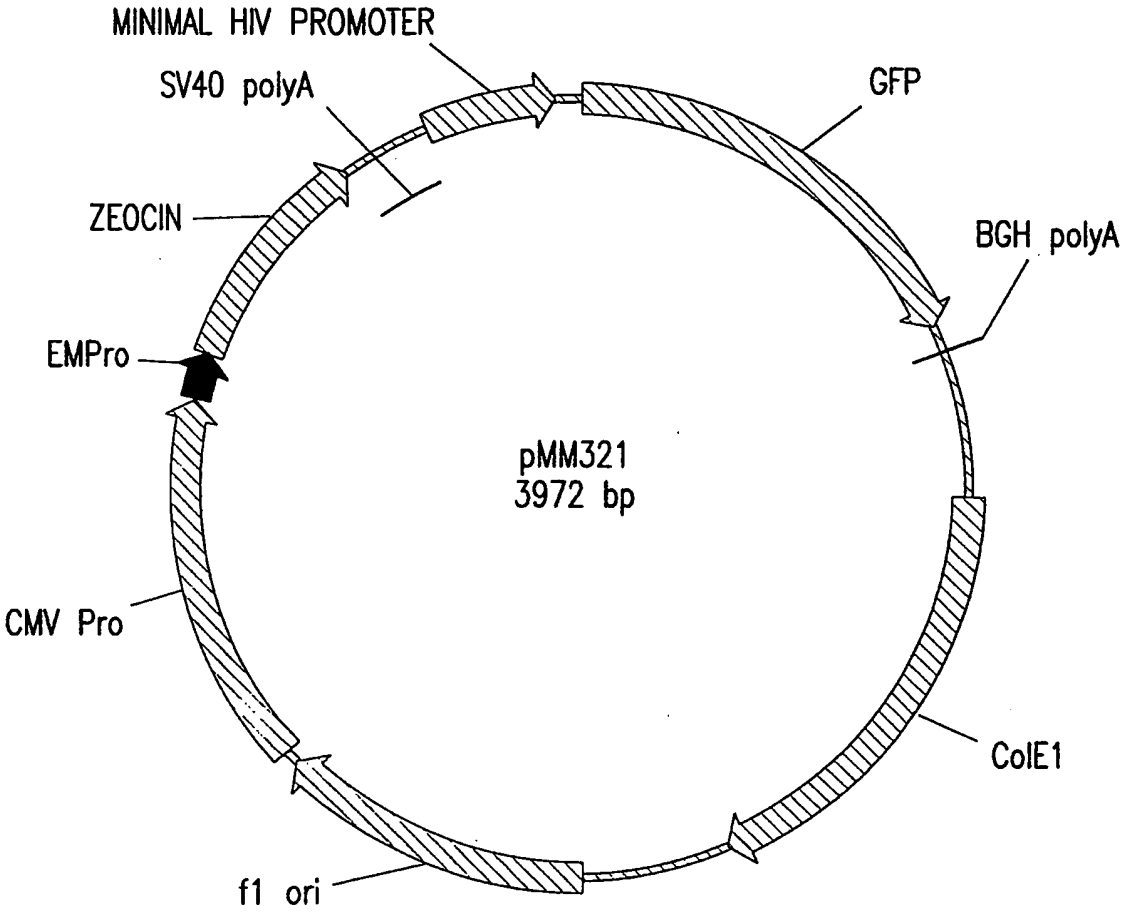

MINIMAL HIV PROMOTER
SV40 polyA
GFP
ZEOCIN
BGH polyA
EMPro
pMM321
3972 bp
CMV Pro
ColE1
f1 ori

FIG.25A

(SEQ ID NO: 19 AND 20)

```
   1 ATGGTGAGCA AGGGCGAGGA GCTGTTCACC GGGGTGGTGC CCATCCTGGT
     TACCACTCGT TCCCGCTCCT CGACAAGTGG CCCCACCACG GGTAGGACCA
  51 CGAGCTGGAC GGCGACGTAA ACGGCCACAA GTTCAGCGTG TCCGGCGAGG
     GCTCGACCTG CCGCTGCATT TGCCGGTGTT CAAGTCGCAC AGGCCGCTCC
 101 GCGAGGGCGA TGCCACCTAC GGCAAGCTGA CCCTGAAGTT CATCTGCACC
     CGCTCCCGCT ACGGTGGATG CCGTTCGACT GGGACTTCAA GTAGACGTGG
 151 ACCGGCAAGC TGCCCGTGCC CTGGCCCACC CTCGTGACCA CCTTCACCTA
     TGGCCGTTCG ACGGGCACGG GACCGGGTGG GAGCACTGGT GGAAGTGGAT
 201 CGGCGTGCAG TGCTTCGCCC GCTACCCCGA CCACATGAAG CAGCACGACT
     GCCGCACGTC ACGAAGCGGG CGATGGGGCT GGTGTACTTC GTCGTGCTGA
 251 TCTTCAAGTC CGCCATGCCC GAAGGCTACG TCCAGGAGCG CACCATCTTC
     AGAAGTTCAG GCGGTACGGG CTTCCGATGC AGGTCCTCGC GTGGTAGAAG
 301 TTCAAGGACG ACGGCAACTA CAAGACCCGC GCCGAGGTGA AGTTCGAGGG
     AAGTTCCTGC TGCCGTTGAT GTTCTGGGCG CGGCTCCACT TCAAGCTCCC
 351 CGACACCCTG GTGAACCGCA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG
     GCTGTGGGAC CACTTGGCGT AGCTCGACTT CCCGTAGCTG AAGTTCCTCC
 401 ACGGCAACAT CCTGGGGCAC AAGCTGGAGT ACAACTACAA CAGCCACAAG
     TGCCGTTGTA GGACCCCGTG TTCGACCTCA TGTTGATGTT GTCGGTGTTC
 451 GTCTATATCA CCGCCGACAA GCAGAAGAAC GGCATCAAGG TGAACTTCAA
     CAGATATAGT GGCGGCTGTT CGTCTTCTTG CCGTAGTTCC ACTTGAAGTT
 501 GACCCGCCAC AACATCGAGG ACGGCAGCGT GCAGCTCGCC GACCACTACC
     CTGGGCGGTG TTGTAGCTCC TGCCGTCGCA CGTCGAGCGG CTGGTGATGG
 551 AGCAGAACAC CCCCATCGGC GACGGCCCCG TGCTGCTGCC CGACAACCAC
     TCGTCTTGTG GGGGTAGCCG CTGCCGGGGC ACGACGACGG GCTGTTGGTG
 601 TACCTGAGCA CCCAGTCCGC CCTGAGCAAA GACCCCAACG AGAAGCGCGA
     ATGGACTCGT GGGTCAGGCG GGACTCGTTT CTGGGGTTGC TCTTCGCGCT
 651 TCACATGGTC CTGCTGGAGT TCGTGACCGC CGCCGGGATC ACTCTCGGCA
     AGTGTACCAG GACGACCTCA AGCACTGGCG GCGGCCCTAG TGAGAGCCGT
 701 TGGACGAGCT GTACAAGTAA CTCGAGTCTA GAGGGCCCGT TTAAACCCGC
     ACCTGCTCGA CATGTTCATT GAGCTCAGAT CTCCCGGGCA AATTTGGGCG
 751 TGATCAGCCT CGACTGTGCC TTCTAGTTGC CAGCCATCTG TTGTTTGCCC
     ACTAGTCGGA GCTGACACGG AAGATCAACG GTCGGTAGAC AACAAACGGG
 801 CTCCCCCGTG CCTTCCTTGA CCCTGGAAGG TGCCACTCCC ACTGTCCTTT
     GAGGGGGCAC GGAAGGAACT GGGACCTTCC ACGGTGAGGG TGACAGGAAA
 851 CCTAATAAAA TGAGGAAATT GCATCGCATT GTCTGAGTAG GTGTCATTCT
     GGATTATTTT ACTCCTTTAA CGTAGCGTAA CAGACTCATC CACAGTAAGA
 901 ATTCTGGGGG GTGGGGTGGG GCAGGACAGC AAGGGGGAGG ATTGGGAAGA
     TAAGACCCCC CACCCCACCC CGTCCTGTCG TTCCCCCTCC TAACCCTTCT
 951 CAATAGCAGG CATGCTGGGG ATGCGGTGGG CTCTATGGCT TCTGAGGCGG
     GTTATCGTCC GTACGACCCC TACGCCACCC GAGATACCGA AGACTCCGCC
1001 AAAGAACCAG CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA
     TTTCTTGGTC GTACACTCGT TTTCCGGTCG TTTTCCGGTC CTTGGCATTT
1051 AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC CTGACGAGCA
     TTCCGGCGCA ACGACCGCAA AAAGGTATCC GAGGCGGGGG GACTGCTCGT
1101 TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
     AGTGTTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC TGTCCTGATA
1151 AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT
     TTTCTATGGT CCGCAAAGGG GGACCTTCGA GGGAGCACGC GAGAGGACAA
1201 CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCGGGAAG
     GGCTGGGACG GCGAATGGCC TATGGACAGG CGGAAAGAGG GAAGCCCTTC
1251 CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
     GCACCGCGAA AGAGTATCGA GTGCGACATC CATAGAGTCA AGCCACATCC
1301 TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC
     AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA AGTCGGGCTG
1351 CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA
     CCGACGCGGA ATAGGCCATT GATAGCAGAA CTCAGGTTGG GCCATTCTGT
1401 CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA
     GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCCTAA TCGTCTCGCT
```

FIG.25B

```
1451 GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACTACGGC
     CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG ATTGATGCCG
1501 TACACTAGAA GAACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC
     ATGTGATCTT CTTGTCATAA ACCATAGACG CGAGACGACT TCGGTCAATG
1551 CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG
     GAAGCCTTTT TCTCAACCAT CGAGAACTAG GCCGTTTGTT TGGTGGCGAC
1601 GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA
     CATCGCCACC AAAAAAACAA ACGTTCGTCG TCTAATGCGC GTCTTTTTTT
1651 GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG ACGCTCAGTG
     CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC TGCGAGTCAC
1701 GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGACATTA ACCTATAAAA
     CTTGCTTTTG AGTGCAATTC CCTAAAACCA GTACTGTAAT TGGATATTTT
1751 ATAGGCGTAT CACGAGGCCC TTTCGTCTCG CGCGTTTCGG TGATGACGGT
     TATCCGCATA GTGCTCCGGG AAAGCAGAGC GCGCAAAGCC ACTACTGCCA
1801 GAAAACCTCT GACACATGCA GCTCCCGGAG ACGGTCACAG CTTGTCTGTA
     CTTTTGGAGA CTGTGTACGT CGAGGGCCTC TGCCAGTGTC GAACAGACAT
1851 AGCGGATGCC GGGAGCAGAC AAGCCCGTCA GGGCGCGTCA GCGGGTGTTG
     TCGCCTACGG CCCTCGTCTG TTCGGGCAGT CCCGCGCAGT CGCCCACAAC
1901 GCGGGTGTCG GGGCTGGCTT AACTATGCGG CATCAGAGCA GATTGTACTG
     CGCCCACAGC CCCGACCGAA TTGATACGCC GTAGTCTCGT CTAACATGAC
1951 AGAGTGCACC ATATGCGGTG TGAAATACCG CACAGATGCG TAAGGAGAAA
     TCTCACGTGG TATACGCCAC ACTTTATGGC GTGTCTACGC ATTCCTCTTT
2001 ATACCGCATC AGGACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT
     TATGGCGTAG TCCTGCGCGG GACATCGCCG CGTAATTCGC GCCGCCCACA
2051 GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG
     CCACCAATGC GCGTCGCACT GGCGATGTGA ACGGTCGCGG GATCGCGGGC
2101 CTCCTTTCGC TTTCTTCCCT TCCTTTCTCG CCACGTTCGC CGGCTTTCCC
     GAGGAAAGCG AAAGAAGGGA AGGAAAGAGC GGTGCAAGCG GCCGAAAGGG
2151 CGTCAAGCTC TAAATCGGGG GCTCCCTTTA GGGTTCCGAT TTAGTGCTTT
     GCAGTTCGAG ATTTAGCCCC CGAGGGAAAT CCCAAGGCTA AATCACGAAA
2201 ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTGATGGT TCACGTAGTG
     TGCCGTGGAG CTGGGGTTTT TTGAACTAAT CCCACTACCA AGTGCATCAC
2251 GGCCATCGCC CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG
     CCGGTAGCGG GACTATCTGC CAAAAAGCGG GAAACTGCAA CCTCAGGTGC
2301 TTCTTTAATA GTGGACTCTT GTTCCAAACT GGAACAACAC TCAACCCTAT
     AAGAAATTAT CACCTGAGAA CAAGGTTTGA CCTTGTTGTG AGTTGGGATA
2351 CTCGGTCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT TCGGCCTATT
     GAGCCAGATA AGAAAACTAA ATATTCCCTA AAACGGCTAA AGCCGGATAA
2401 GGTTAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAACAAA
     CCAATTTTTT ACTCGACTAA ATTGTTTTTA AATTGCGCTT AAAATTGTTT
2451 ATATTAACGC TTACAATTTC CATTCGCCAT TCAGGCTGAA CTAGATCTAG
     TATAATTGCG AATGTTAAAG GTAAGCGGTA AGTCCGACTT GATCTAGATC
2501 AGTCCGTTAC ATAACTTACG GTAAATGGCC CGCCTGGCTG ACCGCCCAAC
     TCAGGCAATG TATTGAATGC CATTTACCGG GCGGACCGAC TGGCGGGTTG
2551 GACCCCCGCC CATTGACGTC AATAATGACG TATGTTCCCA TAGTAACGCC
     CTGGGGGCGG GTAACTGCAG TTATTACTGC ATACAAGGGT ATCATTGCGG
2601 AATAGGGACT TTCCATTGAC GTCAATGGGT GGAGTATTTA CGGTAAACTG
     TTATCCCTGA AAGGTAACTG CAGTTACCCA CCTCATAAAT GCCATTTGAC
2651 CCCACTTGGC AGTACATCAA GTGTATCATA TGCCAAGTAC GCCCCCTATT
     GGGTGAACCG TCATGTAGTT CACATAGTAT ACGGTTCATG CGGGGGATAA
2701 GACGTCAATG ACGGTAAATG GCCCGCCTGG CATTATGCCC AGTACATGAC
     CTGCAGTTAC TGCCATTTAC CGGGCGGACC GTAATACGGG TCATGTACTG
2751 CTTATGGGAC TTTCCTACTT GGCAGTACAT CTACGTATTA GTCATCGCTA
     GAATACCCTG AAAGGATGAA CCGTCATGTA GATGCATAAT CAGTAGCGAT
2801 TTACCATGGT GATGCGGTTT TGGCAGTACA TCAATGGGCG TGGATAGCGG
     AATGGTACCA CTACGCCAAA ACCGTCATGT AGTTACCCGC ACCTATCGCC
2851 TTTGACTCAC GGGGATTTCC AAGTCTCCAC CCCATTGACG TCAATGGGAG
     AAACTGAGTG CCCCTAAAGG TTCAGAGGTG GGGTAACTGC AGTTACCCTC
```

FIG.25C

```
2901 TTTGTTTTGG CACCAAAATC AACGGGACTT TCCAAAATGT CGTAACAACT
     AAACAAAACC GTGGTTTTAG TTGCCCTGAA AGGTTTTACA GCATTGTTGA
2951 CCGCCCCATT GACGCAAATG GGCGGTAGGC GTGTACGGTG GGAGGTCTAT
     GGCGGGGTAA CTGCGTTTAC CCGCCATCCG CACATGCCAC CCTCCAGATA
3001 ATAAGCAGAG CTCGTTTAGT GAACCGTCAG ATCGCCTGGA GACGCCATCC
     TATTCGTCTC GAGCAAATCA CTTGGCAGTC TAGCGGACCT CTGCGGTAGG
3051 ACGCTGTTTT GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGCG
     TGCGACAAAA CTGGAGGTAT CTTCTGTGGC CCTGGCTAGG TCGGAGGCGC
3101 GCCGGGAACG GTGCATTGGA ACGGACCGTG TTGACAATTA ATCATCGGCA
     CGGCCCTTGC CACGTAACCT TGCCTGGCAC AACTGTTAAT TAGTAGCCGT
3151 TAGTATATCG GCATAGTATA ATACGACAAG GTGAGGAACT AAACCATGGC
     ATCATATAGC CGTATCATAT TATGCTGTTC CACTCCTTGA TTTGGTACCG
3201 CAAGTTGACC AGTGCCGTTC CGGTGCTCAC CGCGCGCGAC GTCGCCGGAG
     GTTCAACTGG TCACGGCAAG GCCACGAGTG GCGCGCGCTG CAGCGGCCTC
3251 CGGTCGAGTT CTGGACCGAC CGGCTCGGGT TCTCCCGGGA CTTCGTGGAG
     GCCAGCTCAA GACCTGGCTG GCCGAGCCCA AGAGGGCCCT GAAGCACCTC
3301 GACGACTTCG CCGGTGTGGT CCGGGACGAC GTGACCCTGT TCATCAGCGC
     CTGCTGAAGC GGCCACACCA GGCCCTGCTG CACTGGGACA AGTAGTCGCG
3351 GGTCCAGGAC CAGGTGGTGC CGGACAACAC CCTGGCCTGG GTGTGGGTGC
     CCAGGTCCTG GTCCACCACG GCCTGTTGTG GGACCGGACC CACACCCACG
3401 GCGGCCTGGA CGAGCTGTAC GCCGAGTGGT CGGAGGTCGT GTCCACGAAC
     CGCCGGACCT GCTCGACATG CGGCTCACCA GCCTCCAGCA CAGGTGCTTG
3451 TTCCGGGACG CCTCCGGGCC GGCCATGACC GAGATCGGCG AGCAGCCGTG
     AAGGCCCTGC GGAGGCCCGG CCGGTACTGG CTCTAGCCGC TCGTCGGCAC
3501 GGGGCGGGAG TTCGCCCTGC GCGACCCGGC CGGCAACTGC GTGCACTTCG
     CCCCGCCCTC AAGCGGGACG CGCTGGGCCG GCCGTTGACG CACGTGAAGC
3551 TGGCCGAGGA GCAGGACTGA CACTCGACCT CGAAACTTGT TTATTGCAGC
     ACCGGCTCCT CGTCCTGACT GTGAGCTGGA GCTTTGAACA AATAACGTCG
3601 TTATAATGGT TACAAATAAA GCAATAGCAT CACAAATTTC ACAAATAAAG
     AATATTACCA ATGTTTATTT CGTTATCGTA GTGTTTAAAG TGTTTATTTC
3651 CATTTTTTTC ACTGCATTCT AGTTGTGGTT TGTCCAAACT CATCAATGTA
     GTAAAAAAAG TGACGTAAGA TCAACACCAA ACAGGTTTGA GTAGTTACAT
3701 TCTTATCATG TCTGGATCGA TACTTCAAGA ACTGCTGACA TCGAGCTTGC
     AGAATAGTAC AGACCTAGCT ATGAAGTTCT TGACGACTGT AGCTCGAACG
3751 TACAAGGGAC TTTCCGCTGG GGACTTTCCA GGGAGGCGTG GCCTGGGCGG
     ATGTTCCCTG AAAGGCGACC CCTGAAAGGT CCCTCCGCAC CGGACCCGCC
3801 GACTGGGGAG TGGCGAGCCC TCAGATCCTG CATATAAGCA GCTGCTTTTT
     CTGACCCCTC ACCGCTCGGG AGTCTAGGAC GTATATTCGT CGACGAAAAA
3851 GCCTGTACTG GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC
     CGGACATGAC CCAGAGAGAC CAATCTGGTC TAGACTCGGA CCCTCGAGAG
3901 TGGCTAACTA GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGGTACCGA
     ACCGATTGAT CCCTTGGGTG ACGAATTCGG AGTTATTTCG AACCATGGCT
3951 GCTCGGATCC GAATTCGCCA CC
     CGAGCCTAGG CTTAAGCGGT GG
```

FIG.25D

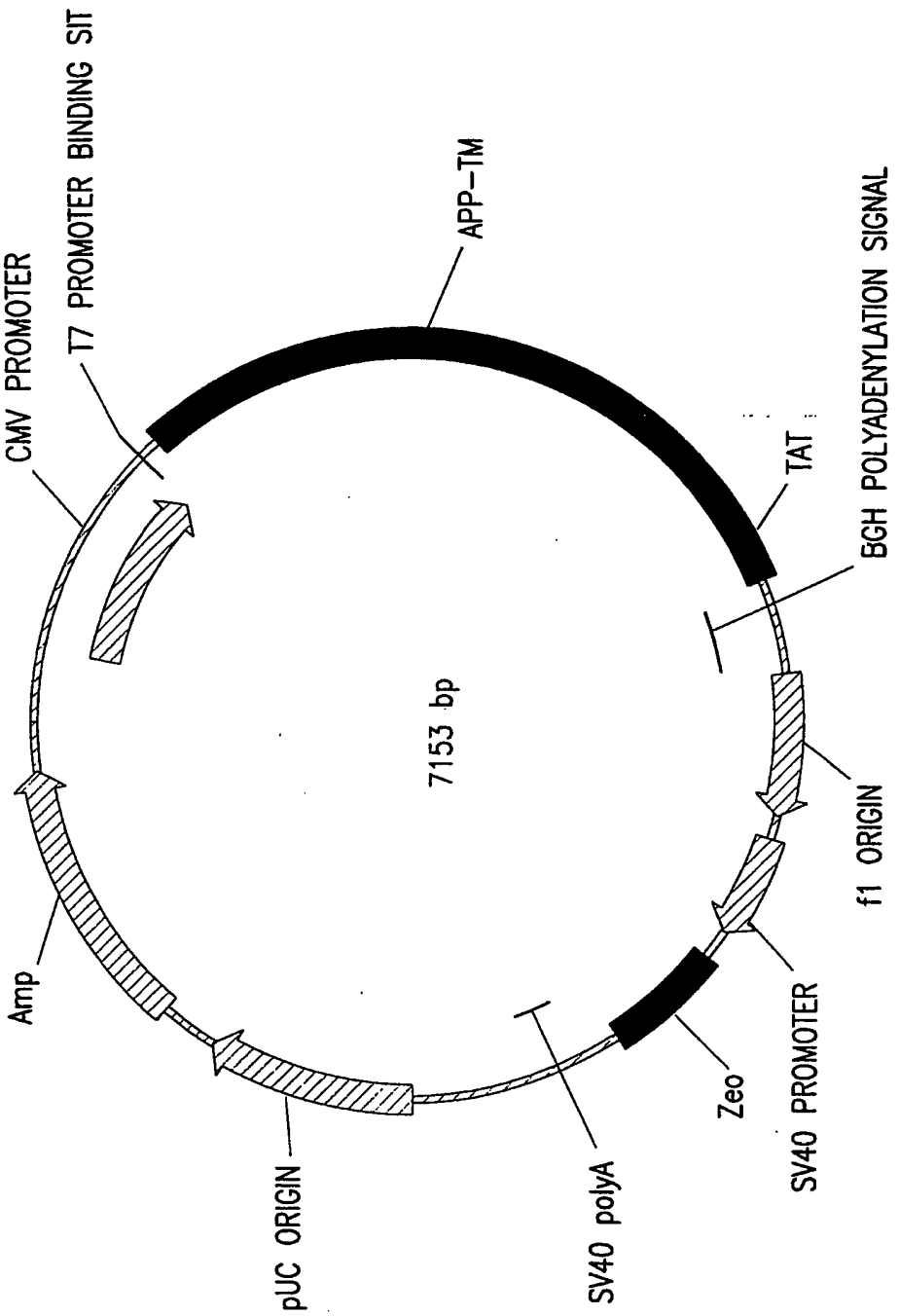

CMV PROMOTER
T7 PROMOTER BINDING SIT
APP-TM
TAT
BGH POLYADENYLATION SIGNAL
7153 bp
f1 ORIGIN
SV40 PROMOTER
Zeo
SV40 polyA
pUC ORIGIN
Amp

FIG.26A

(SEQ ID NO: 21 AND 22)

```
   1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC
     CTGCCTAGCC CTCTAGAGGG CTAGGGGATA CCAGCTGAGA GTCATGTTAG
  51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT
     ACGAGACTAC GGCGTATCAA TTCGGTCATA GACGAGGGAC GAACACACAA
 101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT TTAAGCTACA ACAAGGCAAG
     CCTCCAGCGA CTCATCACGC GCTCGTTTTA AATTCGATGT TGTTCCGTTC
 151 GCTTGACCGA CAATTGCATG AAGAATCTGC TTAGGGTTAG GCGTTTTGCG
     CGAACTGGCT GTTAACGTAC TTCTTAGACG AATCCCAATC CGCAAAACGC
 201 CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT GATTATTGAC
     GACGAAGCGC TACATGCCCG GTCTATATGC GCAACTGTAA CTAATAACTG
 251 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
     ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT
 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG
     ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC
 351 CCCAACGACC CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT
     GGGTTGCTGG GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA
 401 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAC TATTTACGGT
     TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTG ATAAATGCCA
 451 AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC
     TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG
 501 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA
     GGATAACTGC AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT
 551 CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
     GTACTGGAAT ACCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT
 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
     AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 651 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA
     ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT
 701 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA
     ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT
 751 ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG
     TGTTGAGGCG GGGTAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC
 801 GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA CTGCTTACTG
     CAGATATATT CGTCTCGAGA GACCGATTGA TCTCTTGGGT GACGAATGAC
 851 GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
     CGAATAGCTT TAATTATGCT GAGTGATATC CCTCTGGGTT CGACCGATCG
 901 GTTTAAACTT AAGCTTCCCC GCGCAGGGTC GCGATGCTGC CCGGTTTGGC
     CAAATTTGAA TTCGAAGGGG CGCGTCCCAG CGCTACGACG GGCCAAACCG
 951 ACTGCTCCTG CTGGCCGCCT GGACGGCTCG GGCGCTGGAG GTACCCACTG
     TGACGAGGAC GACCGGCGGA CCTGCCGAGC CCGCGACCTC CATGGGTGAC
1001 ATGGTAATGC TGGCCTGCTG GCTGAACCCC AGATTGCCAT GTTCTGTGGC
     TACCATTACG ACCGGACGAC CGACTTGGGG TCTAACGGTA CAAGACACCG
1051 AGACTGAACA TGCACATGAA TGTCCAGAAT GGGAAGTGGG ATTCAGATCC
     TCTGACTTGT ACGTGTACTT ACAGGTCTTA CCCTTCACCC TAAGTCTAGG
1101 ATCAGGGACC AAAACCTGCA TTGATACCAA GGAAGGCATC CTGCAGTATT
     TAGTCCCTGG TTTTGGACGT AACTATGGTT CCTTCCGTAG GACGTCATAA
1151 GCCAAGAAGT CTACCCTGAA CTGCAGATCA CCAATGTGGT AGAAGCCAAC
     CGGTTCTTCA GATGGGACTT GACGTCTAGT GGTTACACCA TCTTCGGTTG
1201 CAACCAGTGA CCATCCAGAA CTGGTGCAAG CGGGGCCGCA AGCAGTGCAA
     GTTGGTCACT GGTAGGTCTT GACCACGTTC GCCCCGGCGT TCGTCACGTT
1251 GACCCATCCC CACTTTGTGA TTCCCTACCG CTGCTTAGTT GGTGAGTTTA
     CTGGGTAGGG GTGAAACACT AAGGGATGGC GACGAATCAA CCACTCAAAT
1301 TAAGTGATGC CCTTCTCGTT CCTGACAAGT GCAAATTCTT ACACCAGGAG
     ATTCACTACG GGAAGAGCAA GGACTGTTCA CGTTTAAGAA TGTGGTCCTC
1351 AGGATGGATG TTTGCGAAAC TCATCTTCAC TGGCACACCG TCGCCAAAGA
     TCCTACCTAC AAACGCTTTG AGTAGAAGTG ACCGTGTGGC AGCGGTTTCT
```

FIG.26B

```
1401 GACATGCAGT GAGAAGAGTA CCAACTTGCA TGACTACGGC ATGTTGCTGC
     CTGTACGTCA CTCTTCTCAT GGTTGAACGT ACTGATGCCG TACAACGACG
1451 CCTGCGGAAT TGACAAGTTC CGAGGGGTAG AGTTTGTGTG TTGCCCACTG
     GGACGCCTTA ACTGTTCAAG GCTCCCCATC TCAAACACAC AACGGGTGAC
1501 GCTGAAGAAA GTGACAATGT GGATTCTGCT GATGCGGAGG AGGATGACTC
     CGACTTCTTT CACTGTTACA CCTAAGACGA CTACGCCTCC TCCTACTGAG
1551 GGATGTCTGG TGGGGCGGAG CAGACACAGA CTATGCAGAT GGGAGTGAAG
     CCTACAGACC ACCCCGCCTC GTCTGTGTCT GATACGTCTA CCCTCACTTC
1601 ACAAAGTAGT AGAAGTAGCA GAGGAGGAAG AAGTGGCTGA GGTGGAAGAA
     TGTTTCATCA TCTTCATCGT CTCCTCCTTC TTCACCGACT CCACCTTCTT
1651 GAAGAAGCCG ATGATGACGA GGACGATGAG GATGGTGATG AGGTAGAGGA
     CTTCTTCGGC TACTACTGCT CCTGCTACTC CTACCACTAC TCCATCTCCT
1701 AGAGGCTGAG GAACCCTACG AAGAAGCCAC AGAGAGAACC ACCAGCATTG
     TCTCCGACTC CTTGGGATGC TTCTTCGGTG TCTCTCTTGG TGGTCGTAAC
1751 CCACCACCAC CACCACCACC ACAGAGTCTG TGGAAGAGGT GGTTCGAGTT
     GGTGGTGGTG GTGGTGGTGG TGTCTCAGAC ACCTTCTCCA CCAAGCTCAA
1801 CCTACAACAG CAGCCAGTAC CCCTGATGCC GTTGACAAGT ATCTCGAGAC
     GGATGTTGTC GTCGGTCATG GGGACTACGG CAACTGTTCA TAGAGCTCTG
1851 ACCTGGGGAT GAGAATGAAC ATGCCCATTT CCAGAAAGCC AAAGAGAGGC
     TGGACCCCTA CTCTTACTTG TACGGGTAAA GGTCTTTCGG TTTCTCTCCG
1901 TTGAGGCCAA GCACCGAGAG AGAATGTCCC AGGTCATGAG AGAATGGGAA
     AACTCCGGTT CGTGGCTCTC TCTTACAGGG TCCAGTACTC TCTTACCCTT
1951 GAGGCAGAAC GTCAAGCAAA GAACTTGCCT AAAGCTGATA AGAAGGCAGT
     CTCCGTCTTG CAGTTCGTTT CTTGAACGGA TTTCGACTAT TCTTCCGTCA
2001 TATCCAGCAT TTCCAGGAGA AAGTGGAATC TTTGGAACAG GAAGCAGCCA
     ATAGGTCGTA AAGGTCCTCT TTCACCTTAG AAACCTTGTC CTTCGTCGGT
2051 ACGAGAGACA GCAGCTGGTG GAGACACACA TGGCCAGAGT GGAAGCCATG
     TGCTCTCTGT CGTCGACCAC CTCTGTGTGT ACCGGTCTCA CCTTCGGTAC
2101 CTCAATGACC GCCGCCGCCT GGCCCTGGAG AACTACATCA CCGCTCTGCA
     GAGTTACTGG CGGCGGCGGA CCGGGACCTC TTGATGTAGT GGCGAGACGT
2151 GGCTGTTCCT CCTCGGCCTC GTCACGTGTT CAATATGCTA AAGAAGTATG
     CCGACAAGGA GGAGCCGGAG CAGTGCACAA GTTATACGAT TTCTTCATAC
2201 TCCGCGCAGA ACAGAAGGAC AGACAGCACA CCCTAAAGCA TTTCGAGCAT
     AGGCGCGTCT TGTCTTCCTG TCTGTCGTGT GGGATTTCGT AAAGCTCGTA
2251 GTGCGCATGG TGGATCCCAA GAAAGCCGCT CAGATCCGGT CCCAGGTTAT
     CACGCGTACC ACCTAGGGTT CTTTCGGCGA GTCTAGGCCA GGGTCCAATA
2301 GACACACCTC CGTGTGATTT ATGAGCGCAT GAATCAGTCT CTCTCCCTGC
     CTGTGTGGAG GCACACTAAA TACTCGCGTA CTTAGTCAGA GAGAGGGACG
2351 TCTACAACGT GCCTGCAGTG GCCGAGGAGA TTCAGGATGA AGTTGATGAG
     AGATGTTGCA CGGACGTCAC CGGCTCCTCT AAGTCCTACT TCAACTACTC
2401 CTGCTTCAGA AAGAGCAAAA CTATTCAGAT GACGTCTTGG CCAACATGAT
     GACGAAGTCT TTCTCGTTTT GATAAGTCTA CTGCAGAACC GGTTGTACTA
2451 TAGTGAACCA AGGATCAGTT ACGGAAACGA TGCTCTCATG CCATCTTTGA
     ATCACTTGGT TCCTAGTCAA TGCCTTTGCT ACGAGAGTAC GGTAGAAACT
2501 CCGAAACGAA AACCACCGTG GAGCTCCTTC CCGTGAATGG AGAGTTCAGC
     GGCTTTGCTT TTGGTGGCAC CTCGAGGAAG GGCACTTACC TCTCAAGTCG
2551 CTGGACGATC TCCAGCCGTG GCATTCTTTT GGGGCTGACT CTGTGCCAGC
     GACCTGCTAG AGGTCGGCAC CGTAAGAAAA CCCCGACTGA GACACGGTCG
2601 CAACACAGAA AACGAAGTTG AGCCTGTTGA TGCCCGCCCT GCTGCCGACC
     GTTGTGTCTT TTGCTTCAAC TCGGACAACT ACGGGCGGGA CGACGGCTGG
2651 GAGGACTGAC CACTCGACCA GGTTCTGGGT TGACAAATAT CAAGACGGAG
     CTCCTGACTG GTGAGCTGGT CCAAGACCCA ACTGTTTATA GTTCTGCCTC
2701 GAGATCTCTG AAGTGAATCT AGATGCAGAA TTCCGACATG ACTCAGGATA
     CTCTAGAGAC TTCACTTAGA TCTACGTCTT AAGGCTGTAC TGAGTCCTAT
2751 TGAAGTTCAT CATCAAAAAT TGGTGTTCTT TGCAGAAGAT GTGGGTTCAA
     ACTTCAAGTA GTAGTTTTTA ACCACAAGAA ACGTCTTCTA CACCCAAGTT
2801 ACAAAGGTGC AATCATTGGA CTCATGGTGG GCGGTGTTGT CATAGCGACA
     TGTTTCCACG TTAGTAACCT GAGTACCACC CGCCACAACA GTATCGCTGT
```

FIG.26C

```
2851 GTGATCGTCA TCACCTTGGT GATGCTGAAG AAGAAAGATA TCATGGAGCC
     CACTAGCAGT AGTGGAACCA CTACGACTTC TTCTTTCTAT AGTACCTCGG
2901 AGTAGATCCT AGACTAGAGC CCTGGAAGCA TCCAGGAAGT CAGCCTAAAA
     TCATCTAGGA TCTGATCTCG GGACCTTCGT AGGTCCTTCA GTCGGATTTT
2951 CTGCTTGTAC CAATTGCTAT TGTAAAAAGT GTTGCTTTCA TTGCCAAGTT
     GACGAACATG GTTAACGATA ACATTTTTCA CAACGAAAGT AACGGTTCAA
3001 TGTTTCATGA CAAAAGCCTT AGGCATCTCC TATGGCAGGA AGAAGCGGAG
     ACAAAGTACT GTTTTCGGAA TCCGTAGAGG ATACCGTCCT TCTTCGCCTC
3051 ACAGCGACGA AGAGCTCATC AGAACAGTCA GACTCATCAA GCTTCTCTAT
     TGTCGCTGCT TCTCGAGTAG TCTTGTCAGT CTGAGTAGTT CGAAGAGATA
3101 CAAAGCAGTA AGTAGGCGGC CGCTCGAGTC TAGAGGGCCC GTTTAAACCC
     GTTTCGTCAT TCATCCGCCG GCGAGCTCAG ATCTCCCGGG CAAATTTGGG
3151 GCTGATCAGC CTCGACTGTG CCTTCTAGTT GCCAGCCATC TGTTGTTTGC
     CGACTAGTCG GAGCTGACAC GGAAGATCAA CGGTCGGTAG ACAACAAACG
3201 CCCTCCCCCG TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT
     GGGAGGGGGC ACGGAAGGAA CTGGGACCTT CCACGGTGAG GGTGACAGGA
3251 TTCCTAATAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT
     AAGGATTATT TTACTCCTTT AACGTAGCGT AACAGACTCA TCCACAGTAA
3301 CTATTCTGGG GGGTGGGGTG GGGCAGGACA GCAAGGGGGA GGATTGGGAA
     GATAAGACCC CCCACCCCAC CCCGTCCTGT CGTTCCCCCT CCTAACCCTT
3351 GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG CTTCTGAGGC
     CTGTTATCGT CCGTACGACC CCTACGCCAC CCGAGATACC GAAGACTCCG
3401 GGAAAGAACC AGCTGGGGCT CTAGGGGGTA TCCCCACGCG CCCTGTAGCG
     CCTTTCTTGG TCGACCCCGA GATCCCCCAT AGGGGTGCGC GGGACATCGC
3451 GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA
     CGCGTAATTC GCGCCGCCCA CACCACCAAT GCGCGTCGCA CTGGCGATGT
3501 CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT
     GAACGGTCGC GGGATCGCGG GCGAGGAAAG CGAAAGAAGG GAAGGAAAGA
3551 CGCCACGTTC GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGCATCCCTT
     GCGGTGCAAG CGGCCGAAAG GGGCAGTTCG AGATTTAGCC CCGTAGGGAA
3601 TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT
     ATCCCAAGGC TAAATCACGA AATGCCGTGG AGCTGGGGTT TTTTGAACTA
3651 TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG
     ATCCCACTAC CAAGTGCATC ACCCGGTAGC GGGACTATCT GCCAAAAAGC
3701 CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA
     GGGAAACTGC AACCTCAGGT GCAAGAAATT ATCACCTGAG AACAAGGTTT
3751 CTGGAACAAC ACTCAACCCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG
     GACCTTGTTG TGAGTTGGGA TAGAGCCAGA TAAGAAAACT AAATATTCCC
3801 ATTTTGGGGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA
     TAAAACCCCT AAAGCCGGAT AACCAATTTT TTACTCGACT AAATTGTTTT
3851 ATTTAACGCG AATTAATTCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA
     TAAATTGCGC TTAATTAAGA CACCTTACAC ACAGTCAATC CCACACCTTT
3901 GTCCCCAGGC TCCCCAGGCA GGCAGAAGTA TGCAAAGCAT GCATCTCAAT
     CAGGGGTCCG AGGGGTCCGT CCGTCTTCAT ACGTTTCGTA CGTAGAGTTA
3951 TAGTCAGCAA CCAGGTGTGG AAAGTCCCCA GGCTCCCCAG CAGGCAGAAG
     ATCAGTCGTT GGTCCACACC TTTCAGGGGT CCGAGGGGTC GTCCGTCTTC
4001 TATGCAAAGC ATGCATCTCA ATTAGTCAGC AACCATAGTC CCGCCCCTAA
     ATACGTTTCG TACGTAGAGT TAATCAGTCG TTGGTATCAG GGCGGGGATT
4051 CTCCGCCCAT CCCGCCCCTA ACTCCGCCCA GTTCCGCCCA TTCTCCGCCC
     GAGGCGGGTA GGGCGGGGAT TGAGGCGGGT CAAGGCGGGT AAGAGGCGGG
4101 CATGGCTGAC TAATTTTTTT TATTTATGCA GAGGCCGAGG CCGCCTCTGC
     GTACCGACTG ATTAAAAAAA ATAAATACGT CTCCGGCTCC GGCGGAGACG
4151 CTCTGAGCTA TTCCAGAAGT AGTGAGGAGG CTTTTTTGGA GGCCTAGGCT
     GAGACTCGAT AAGGTCTTCA TCACTCCTCC GAAAAAACCT CCGGATCCGA
4201 TTTGCAAAAA GCTCCCGGGA GCTTGTATAT CCATTTTCGG ATCTGATCAG
     AAACGTTTTT CGAGGGCCCT CGAACATATA GGTAAAAGCC TAGACTAGTC
```

FIG.26D

```
4251 CACGTGTTGA CAATTAATCA TCGGCATAGT ATATCGGCAT AGTATAATAC
     GTGCACAACT GTTAATTAGT AGCCGTATCA TATAGCCGTA TCATATTATG
4301 GACAAGGTGA GGAACTAAAC CATGGCCAAG TTGACCAGTG CCGTTCCGGT
     CTGTTCCACT CCTTGATTTG GTACCGGTTC AACTGGTCAC GGCAAGGCCA
4351 GCTCACCGCG CGCGACGTCG CCGGAGCGGT CGAGTTCTGG ACCGACCGGC
     CGAGTGGCGC GCGCTGCAGC GGCCTCGCCA GCTCAAGACC TGGCTGGCCG
4401 TCGGGTTCTC CCGGGACTTC GTGGAGGACG ACTTCGCCGG TGTGGTCCGG
     AGCCCAAGAG GGCCCTGAAG CACCTCCTGC TGAAGCGGCC ACACCAGGCC
4451 GACGACGTGA CCCTGTTCAT CAGCGCGGTC CAGGACCAGG TGGTGCCGGA
     CTGCTGCACT GGGACAAGTA GTCGCGCCAG GTCCTGGTCC ACCACGGCCT
4501 CAACACCCTG GCCTGGGTGT GGGTGCGCGG CCTGGACGAG CTGTACGCCG
     GTTGTGGGAC CGGACCCACA CCCACGCGCC GGACCTGCTC GACATGCGGC
4551 AGTGGTCGGA GGTCGTGTCC ACGAACTTCC GGGACGCCTC CGGGCCGGCC
     TCACCAGCCT CCAGCACAGG TGCTTGAAGG CCCTGCGGAG GCCCGGCCGG
4601 ATGACCGAGA TCGGCGAGCA GCCGTGGGGG CGGGAGTTCG CCCTGCGCGA
     TACTGGCTCT AGCCGCTCGT CGGCACCCCC GCCCTCAAGC GGGACGCGCT
4651 CCCGGCCGGC AACTGCGTGC ACTTCGTGGC CGAGGAGCAG GACTGACACG
     GGGCCGGCCG TTGACGCACG TGAAGCACCG GCTCCTCGTC CTGACTGTGC
4701 TGCTACGAGA TTTCGATTCC ACCGCCGCCT TCTATGAAAG GTTGGGCTTC
     ACGATGCTCT AAAGCTAAGG TGGCGGCGGA AGATACTTTC CAACCCGAAG
4751 GGAATCGTTT TCCGGGACGC CGGCTGGATG ATCCTCCAGC GCGGGGATCT
     CCTTAGCAAA AGGCCCTGCG GCCGACCTAC TAGGAGGTCG CGCCCCTAGA
4801 CATGCTGGAG TTCTTCGCCC ACCCCAACTT GTTTATTGCA GCTTATAATG
     GTACGACCTC AAGAAGCGGG TGGGGTTGAA CAAATAACGT CGAATATTAC
4851 GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT
     CAATGTTTAT TTCGTTATCG TAGTGTTTAA AGTGTTTATT TCGTAAAAAA
4901 TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA
     AGTGACGTAA GATCAACACC AAACAGGTTT GAGTAGTTAC ATAGAATAGT
4951 TGTCTGTATA CCGTCGACCT CTAGCTAGAG CTTGGCGTAA TCATGGTCAT
     ACAGACATAT GGCAGCTGGA GATCGATCTC GAACCGCATT AGTACCAGTA
5001 AGCTGTTTCC TGTGTGAAAT TGTTATCCGC TCACAATTCC ACACAACATA
     TCGACAAAGG ACACACTTTA ACAATAGGCG AGTGTTAAGG TGTGTTGTAT
5051 CGAGCCGGAA GCATAAAGTG TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA
     GCTCGGCCTT CGTATTTCAC ATTTCGGACC CCACGGATTA CTCACTCGAT
5101 ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG TCGGGAAACC
     TGAGTGTAAT TAACGCAACG CGAGTGACGG GCGAAAGGTC AGCCCTTTGG
5151 TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT
     ACAGCACGGT CGACGTAATT ACTTAGCCGG TTGCGCGCCC CTCTCCGCCA
5201 TTGCGTATTG GGCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC
     AACGCATAAC CCGCGAGAAG GCGAAGGAGC GAGTGACTGA GCGACGCGAG
5251 GGTCGTTCGG CTGCGGCGAG CGGTATCAGC TCACTCAAAG GCGGTAATAC
     CCAGCAAGCC GACGCCGCTC GCCATAGTCG AGTGAGTTTC CGCCATTATG
5301 GGTTATCCAC AGAATCAGGG GATAACGCAG GAAAGAACAT GTGAGCAAAA
     CCAATAGGTG TCTTAGTCCC CTATTGCGTC CTTTCTTGTA CACTCGTTTT
5351 GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG GCCGCGTTGC TGGCGTTTTT
     CCGGTCGTTT TCCGGTCCTT GGCATTTTTC CGGCGCAACG ACCGCAAAAA
5401 CCATAGGCTC CGCCCCCCTG ACGAGCATCA CAAAAATCGA CGCTCAAGTC
     GGTATCCGAG GCGGGGGGAC TGCTCGTAGT GTTTTTAGCT GCGAGTTCAG
5451 AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC GTTTCCCCCT
     TCTCCACCGC TTTGGGCTGT CCTGATATTT CTATGGTCCG CAAAGGGGGA
5501 GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA
     CCTTCGAGGG AGCACGCGAG AGGACAAGGC TGGGACGGCG AATGGCCTAT
5551 CCTGTCCGCC TTTCTCCCTT CGGGAAGCGT GGCGCTTTCT CAATGCTCAC
     GGACAGGCGG AAAGAGGGAA GCCCTTCGCA CCGCGAAAGA GTTACGAGTG
5601 GCTGTAGGTA TCTCAGTTCG GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT
     CGACATCCAT AGAGTCAAGC CACATCCAGC AAGCGAGGTT CGACCCGACA
```

FIG.26E

```
5651 GTGCACGAAC CCCCCGTTCA GCCCGACCGC TGCGCCTTAT CCGGTAACTA
     CACGTGCTTG GGGGGCAAGT CGGGCTGGCG ACGCGGAATA GGCCATTGAT
5701 TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA CTGGCAGCAG
     AGCAGAACTC AGGTTGGGCC ATTCTGTGCT GAATAGCGGT GACCGTCGTC
5751 CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG
     GGTGACCATT GTCCTAATCG TCTCGCTCCA TACATCCGCC ACGATGTCTC
5801 TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGGA CAGTATTTGG
     AAGAACTTCA CCACCGGATT GATGCCGATG TGATCTTCCT GTCATAAACC
5851 TATCTGCGCT CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT
     ATAGACGCGA GACGACTTCG GTCAATGGAA GCCTTTTTCT CAACCATCGA
5901 CTTGATCCGG CAAACAAACC ACCGCTGGTA GCGGTGGTTT TTTTGTTTGC
     GAACTAGGCC GTTTGTTTGG TGGCGACCAT CGCCACCAAA AAAACAAACG
5951 AAGCAGCAGA TTACGCGCAG AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT
     TTCGTCGTCT AATGCGCGTC TTTTTTTCCT AGAGTTCTTC TAGGAAACTA
6001 CTTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA CGTTAAGGGA
     GAAAAGATGC CCCAGACTGC GAGTCACCTT GCTTTTGAGT GCAATTCCCT
6051 TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT
     AAAACCAGTA CTCTAATAGT TTTTCCTAGA AGTGGATCTA GGAAAATTTA
6101 TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAACTTGGTC
     ATTTTTACTT CAAAATTTAG TTAGATTTCA TATATACTCA TTTGAACCAG
6151 TGACAGTTAC CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC
     ACTGTCAATG GTTACGAATT AGTCACTCCG TGGATAGAGT CGCTAGACAG
6201 TATTTCGTTC ATCCATAGTT GCCTGACTCC CCGTCGTGTA GATAACTACG
     ATAAAGCAAG TAGGTATCAA CGGACTGAGG GGCAGCACAT CTATTGATGC
6251 ATACGGGAGG GCTTACCATC TGGCCCCAGT GCTGCAATGA TACCGCGAGA
     TATGCCCTCC CGAATGGTAG ACCGGGGTCA CGACGTTACT ATGGCGCTCT
6301 CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG CCAGCCGGAA
     GGGTGCGAGT GGCCGAGGTC TAAATAGTCG TTATTTGGTC GGTCGGCCTT
6351 GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT
     CCCGGCTCGC GTCTTCACCA GGACGTTGAA ATAGGCGGAG GTAGGTCAGA
6401 ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT
     TAATTAACAA CGGCCCTTCG ATCTCATTCA TCAAGCGGTC AATTATCAAA
6451 GCGCAACGTT GTTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTCGT
     CGCGTTGCAA CAACGGTAAC GATGTCCGTA GCACCACAGT GCGAGCAGCA
6501 TTGGTATGGC TTCATTCAGC TCCGGTTCCC AACGATCAAG GCGAGTTACA
     AACCATACCG AAGTAAGTCG AGGCCAAGGG TTGCTAGTTC CGCTCAATGT
6551 TGATCCCCCA TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT
     ACTAGGGGGT ACAACACGTT TTTTCGCCAA TCGAGGAAGC CAGGAGGCTA
6601 CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG GTTATGGCAG
     GCAACAGTCT TCATTCAACC GGCGTCACAA TAGTGAGTAC CAATACCGTC
6651 CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG
     GTGACGTATT AAGAGAATGA CAGTACGGTA GGCATTCTAC GAAAAGACAC
6701 ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGGCGACC
     TGACCACTCA TGAGTTGGTT CAGTAAGACT CTTATCACAT ACGCCGCTGG
6751 GAGTTGCTCT TGCCCGGCGT CAATACGGGA TAATACCGCG CCACATAGCA
     CTCAACGAGA ACGGGCCGCA GTTATGCCCT ATTATGGCGC GGTGTATCGT
6801 GAACTTTAAA AGTGCTCATC ATTGGAAAAC GTTCTTCGGG GCGAAAACTC
     CTTGAAATTT TCACGAGTAG TAACCTTTTG CAAGAAGCCC CGCTTTTGAG
6851 TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGTGC
     AGTTCCTAGA ATGGCGACAA CTCTAGGTCA AGCTACATTG GGTGAGCACG
6901 ACCCAACTGA TCTTCAGCAT CTTTTACTTT CACCAGCGTT TCTGGGTGAG
     TGGGTTGACT AGAAGTCGTA GAAAATGAAA GTGGTCGCAA AGACCCACTC
6951 CAAAAACAGG AAGGCAAAAT GCCGCAAAAA AGGGAATAAG GGCGACACGG
     GTTTTTGTCC TTCCGTTTTA CGGCGTTTTT TCCCTTATTC CCGCTGTGCC
7001 AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT GAAGCATTTA
     TTTACAACTT ATGAGTATGA GAAGGAAAAA GTTATAATAA CTTCGTAAAT
```

FIG.26F

```
7051 TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA
     AGTCCCAATA ACAGAGTACT CGCCTATGTA TAAACTTACA TAAATCTTTT
7101 ATAAACAAAT AGGGGTTCCG CGCACATTTC CCCGAAAAGT GCCACCTGAC
     TATTTGTTTA TCCCCAAGGC GCGTGTAAAG GGGCTTTTCA CGGTGGACTG
7151 GTC
     CAG
```

FIG.26G

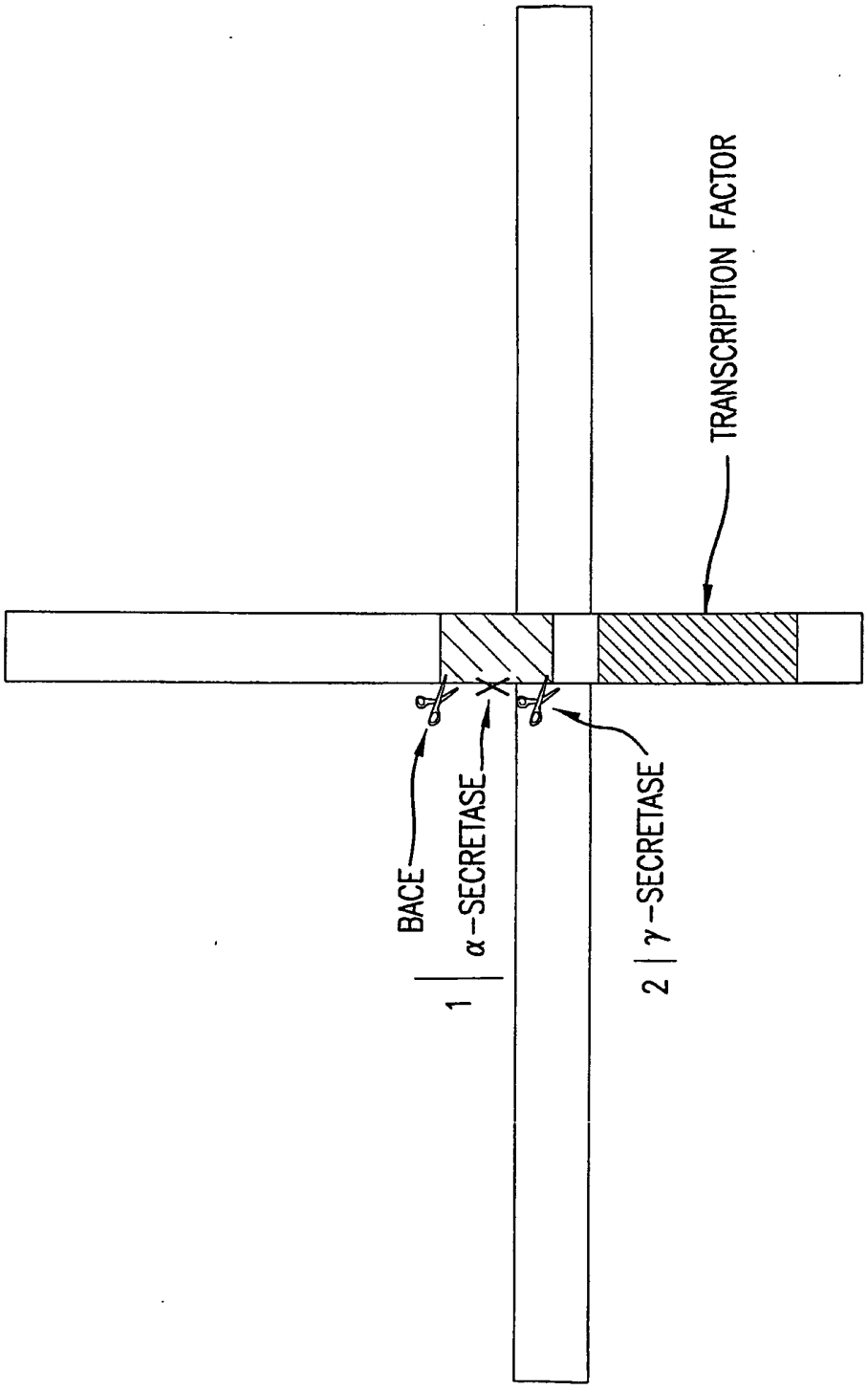


FIG.27A

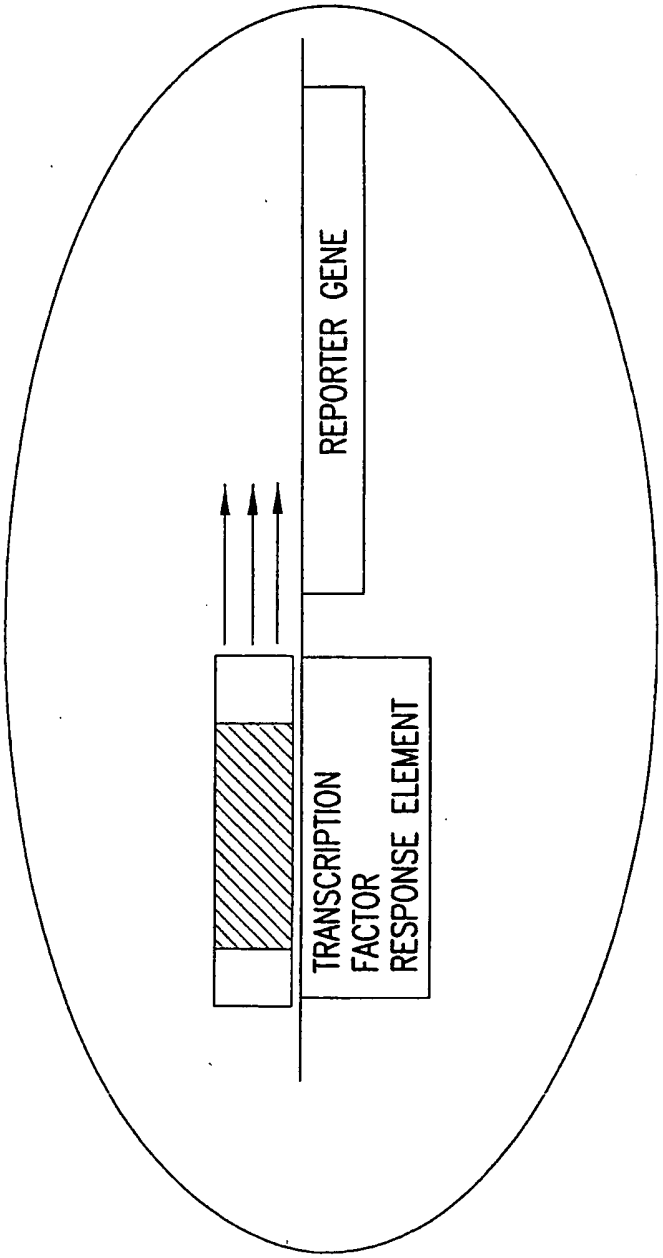


FIG.27B

DNA sequence of APP(1-651)NFEV, K612V-TATexon1(M1L) APP (664-695)
(SEQ ID NO: 23)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC
1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
```

FIG.28A

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAACTTTGA AGTGGAATTC
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC
2051 CAGGAAGTCA GCCTAAAACT GCTTGTACCA ATTGCTATTG TAAAAAGTGT
2101 TGCTTTCATT GCCAAGTTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG.28B

(SEQ ID NO: 24)

Amino acid sequence of APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefslddlqpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

nfevefrhdsgyevhhqvlvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

6

~~adi~~lepvdprlepwkhpgsqpktactncyckkccfhcqvcfmtkalgisygrkkrrqrrrahqnsqthqaslskq

7 8

~~risstvaaa~~daavtpeerhlskmqqngyenptykffeqmqn

FIG.29

DNA sequence of APP(1-651)NFEV, K612V-GAL4VP16(delMet) APP (664-695)

(SEQ ID NO: 25)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
```

FIG.30A

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC

1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC

1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC

1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC

1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG

1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG

1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC

1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA

1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAACTTTGA AGTGGAATTC

1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTTGC

1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG

1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG

1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC

2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTTGCC

2051 GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT

2101 CTGAAGAACA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC

2151 GCTGACTAGG GCACATCTGA CAGAAGTGGA ATCAAGGCTA GAAAGACTGG

2201 AACAGCTATT TCTACTGATT TTTCCTCGAG AAGACCTTGA CATGATTTTG

2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT

2301 ACAAGATAAT GTGAATAAAG ATGCCGTCAC AGATAGATTG GCTTCAGTGG

2351 AGACTGATAT GCCTCTAACA TTGAGACAGC ATAGAATAAG TGCGACATCA

2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG

2451 AATTCCCGGG GATCTGGCCC CCCCGACCGA TGTCAGCCTG GGGGACGAGC

2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC

2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.30B

2601 TACCCCCCAC GACTCCGCCC CCTACGGCGC TCTGGATATG GCCGACTTCG

2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG

2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG

2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT

2801 TCTTTGAGCA GATGCAGAAC TAG

FIG.30C

(SEQ ID NO: 26)

Amino acid sequence of APP(1-651)NFEV, K612V, GAL4-VP16(delM1) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefslddlqpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

nfevefrhdsgyevhhqvlvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

~~adi~~kllssiegacdicrlkklkcskekpkcakclknnwecryspktkrspltrahltevesrlerleqlfllifpredld

6

milkmdslqdikalltglfvqdnvnkdavtdrlasvetdmpltlrqhrisatssseessnkgqrqltvsqipgdlapp

tdvslqdelhldqedvamahadaldqfdldmlqdgdspqpgftphdsapygaldmadfefeqmftdalgidey

7 8

~~ggdiqhsgaaa~~daavtpeerhlskmqqngyenptykffeqmqn

FIG.31

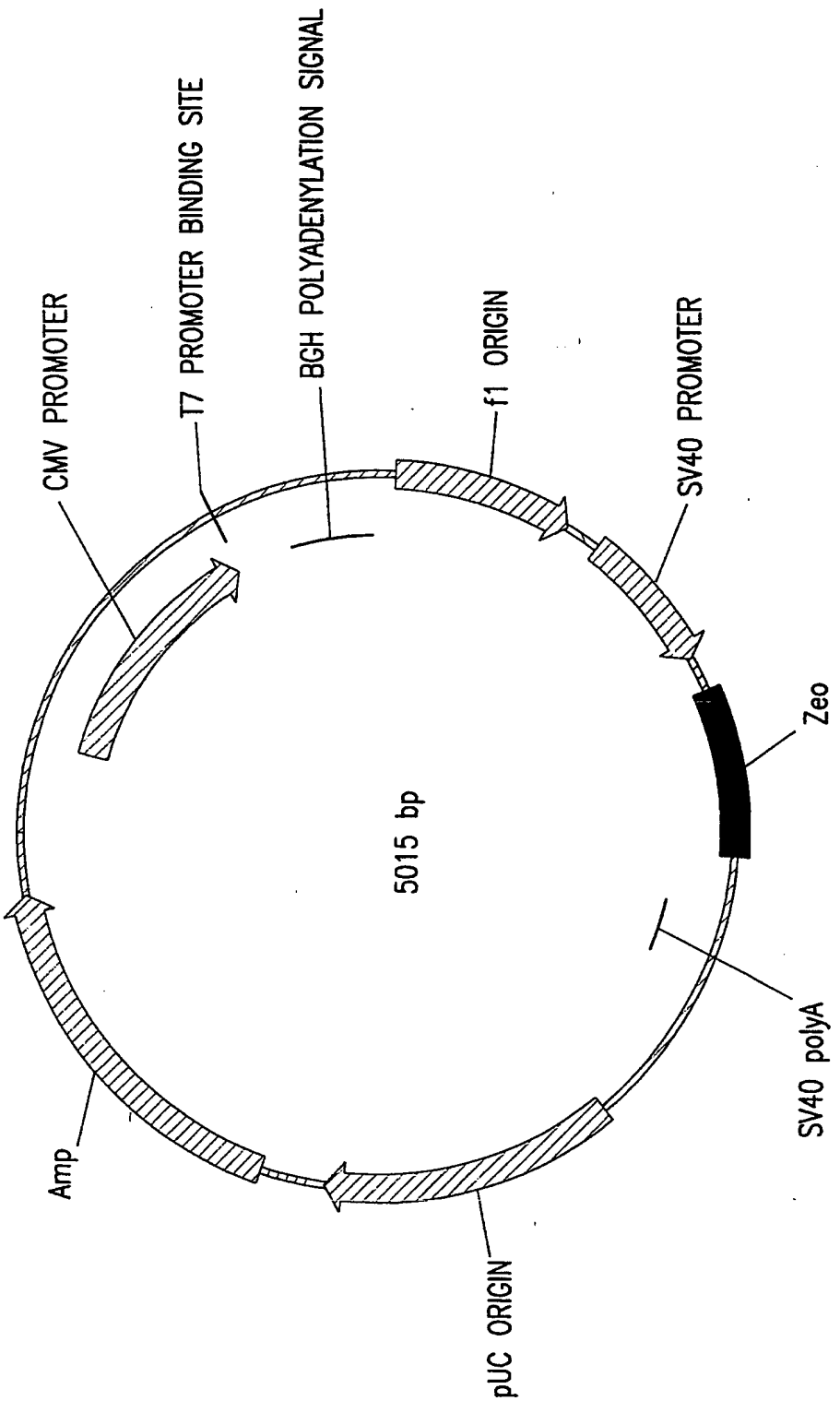

CMV PROMOTER
T7 PROMOTER BINDING SITE
BGH POLYADENYLATION SIGNAL
f1 ORIGIN
SV40 PROMOTER
Zeo
SV40 polyA
pUC ORIGIN
Amp
5015 bp

FIG.32A

(SEQ ID NO: 27 AND 28)

```
   1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC
     CTGCCTAGCC CTCTAGAGGG CTAGGGGATA CCAGCTGAGA GTCATGTTAG
  51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT
     ACGAGACTAC GGCGTATCAA TTCGGTCATA GACGAGGGAC GAACACACAA
 101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT TTAAGCTACA ACAAGGCAAG
     CCTCCAGCGA CTCATCACGC GCTCGTTTTA AATTCGATGT TGTTCCGTTC
 151 GCTTGACCGA CAATTGCATG AAGAATCTGC TTAGGGTTAG GCGTTTTGCG
     CGAACTGGCT GTTAACGTAC TTCTTAGACG AATCCCAATC CGCAAAACGC
 201 CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT GATTATTGAC
     GACGAAGCGC TACATGCCCG GTCTATATGC GCAACTGTAA CTAATAACTG
 251 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
     ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT
 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG
     ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC
 351 CCCAACGACC CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT
     GGGTTGCTGG GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA
 401 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAC TATTTACGGT
     TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTG ATAAATGCCA
 451 AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC
     TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG
 501 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA
     GGATAACTGC AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT
 551 CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
     GTACTGGAAT ACCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT
 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
     AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 651 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA
     ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT
 701 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA
     ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT
 751 ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG
     TGTTGAGGCG GGGTAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC
 801 GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA CTGCTTACTG
     CAGATATATT CGTCTCGAGA GACCGATTGA TCTCTTGGGT GACGAATGAC
 851 GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
     CGAATAGCTT TAATTATGCT GAGTGATATC CCTCTGGGTT CGACCGATCG
 901 GTTTAAACTT AAGCTTGGTA CCGAGCTCGG ATCCACTAGT CCAGTGTGGT
     CAAATTTGAA TTCGAACCAT GGCTCGAGCC TAGGTGATCA GGTCACACCA
 951 GGAATTCTGC AGATATCCAG CACAGTGGCG GCCGCTCGAG TCTAGAGGGC
     CCTTAAGACG TCTATAGGTC GTGTCACCGC CGGCGAGCTC AGATCTCCCG
1001 CCGTTTAAAC CCGCTGATCA GCCTCGACTG TGCCTTCTAG TTGCCAGCCA
     GGCAAATTTG GGCGACTAGT CGGAGCTGAC ACGGAAGATC AACGGTCGGT
1051 TCTGTTGTTT GCCCCTCCCC CGTGCCTTCC TTGACCCTGG AAGGTGCCAC
     AGACAACAAA CGGGGAGGGG GCACGGAAGG AACTGGGACC TTCCACGGTG
1101 TCCCACTGTC CTTTCCTAAT AAAATGAGGA AATTGCATCG CATTGTCTGA
     AGGGTGACAG GAAAGGATTA TTTTACTCCT TTAACGTAGC GTAACAGACT
1151 GTAGGTGTCA TTCTATTCTG GGGGGTGGGG TGGGGCAGGA CAGCAAGGGG
     CATCCACAGT AAGATAAGAC CCCCCACCCC ACCCCGTCCT GTCGTTCCCC
```

FIG.32B

1201 GAGGATTGGG AAGACAATAG CAGGCATGCT GGGGATGCGG TGGGCTCTAT
CTCCTAACCC TTCTGTTATC GTCCGTACGA CCCCTACGCC ACCCGAGATA
1251 GGCTTCTGAG GCGGAAAGAA CCAGCTGGGG CTCTAGGGGG TATCCCCACG
CCGAAGACTC CGCCTTTCTT GGTCGACCCC GAGATCCCCC ATAGGGGTGC
1301 CGCCCTGTAG CGGCGCATTA AGCGCGGCGG GTGTGGTGGT TACGCGCAGC
GCGGGACATC GCCGCGTAAT TCGCGCCGCC CACACCACCA ATGCGCGTCG
1351 GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT
CACTGGCGAT GTGAACGGTC GCGGGATCGC GGGCGAGGAA AGCGAAAGAA
1401 CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC
GGGAAGGAAA GAGCGGTGCA AGCGGCCGAA AGGGGCAGTT CGAGATTTAG
1451 GGGGCATCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC
CCCCGTAGGG AAATCCCAAG GCTAAATCAC GAAATGCCGT GGAGCTGGGG
1501 AAAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA
TTTTTTGAAC TAATCCCACT ACCAAGTGCA TCACCCGGTA GCGGGACTAT
1551 GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC
CTGCCAAAAA GCGGGAAACT GCAACCTCAG GTGCAAGAAA TTATCACCTG
1601 TCTTGTTCCA AACTGGAACA ACACTCAACC CTATCTCGGT CTATTCTTTT
AGAACAAGGT TTGACCTTGT TGTGAGTTGG GATAGAGCCA GATAAGAAAA
1651 GATTTATAAG GGATTTTGGG GATTTCGGCC TATTGGTTAA AAAATGAGCT
CTAAATATTC CCTAAAACCC CTAAAGCCGG ATAACCAATT TTTTACTCGA
1701 GATTTAACAA AAATTTAACG CGAATTAATT CTGTGGAATG TGTGTCAGTT
CTAAATTGTT TTTAAATTGC GCTTAATTAA GACACCTTAC ACACAGTCAA
1751 AGGGTGTGGA AAGTCCCCAG GCTCCCCAGG CAGGCAGAAG TATGCAAAGC
TCCCACACCT TTCAGGGGTC CGAGGGGTCC GTCCGTCTTC ATACGTTTCG
1801 ATGCATCTCA ATTAGTCAGC AACCAGGTGT GGAAAGTCCC CAGGCTCCCC
TACGTAGAGT TAATCAGTCG TTGGTCCACA CCTTTCAGGG GTCCGAGGGG
1851 AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA GCAACCATAG
TCGTCCGTCT TCATACGTTT CGTACGTAGA GTTAATCAGT CGTTGGTATC
1901 TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC CAGTTCCGCC
AGGGCGGGGA TTGAGGCGGG TAGGGCGGGG ATTGAGGCGG GTCAAGGCGG
1951 CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA
GTAAGAGGCG GGGTACCGAC TGATTAAAAA AAATAAATAC GTCTCCGGCT
2001 GGCCGCCTCT GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTTG
CCGGCGGAGA CGGAGACTCG ATAAGGTCTT CATCACTCCT CCGAAAAAAC
2051 GAGGCCTAGG CTTTTGCAAA AAGCTCCCGG GAGCTTGTAT ATCCATTTTC
CTCCGGATCC GAAAACGTTT TTCGAGGGCC CTCGAACATA TAGGTAAAAG
2101 GGATCTGATC AGCACGTGTT GACAATTAAT CATCGGCATA GTATATCGGC
CCTAGACTAG TCGTGCACAA CTGTTAATTA GTAGCCGTAT CATATAGCCG
2151 ATAGTATAAT ACGACAAGGT GAGGAACTAA ACCATGGCCA AGTTGACCAG
TATCATATTA TGCTGTTCCA CTCCTTGATT TGGTACCGGT TCAACTGGTC
2201 TGCCGTTCCG GTGCTCACCG CGCGCGACGT CGCCGGAGCG GTCGAGTTCT
ACGGCAAGGC CACGAGTGGC GCGCGCTGCA GCGGCCTCGC CAGCTCAAGA
2251 GGACCGACCG GCTCGGGTTC TCCCGGGACT TCGTGGAGGA CGACTTCGCC
CCTGGCTGGC CGAGCCCAAG AGGGCCCTGA AGCACCTCCT GCTGAAGCGG
2301 GGTGTGGTCC GGGACGACGT GACCCTGTTC ATCAGCGCGG TCCAGGACCA
CCACACCAGG CCCTGCTGCA CTGGGACAAG TAGTCGCGCC AGGTCCTGGT
2351 GGTGGTGCCG GACAACACCC TGGCCTGGGT GTGGGTGCGC GGCCTGGACG
CCACCACGGC CTGTTGTGGG ACCGGACCCA CACCCACGCG CCGGACCTGC

FIG.32C

```
2401 AGCTGTACGC CGAGTGGTCG GAGGTCGTGT CCACGAACTT CCGGGACGCC
     TCGACATGCG GCTCACCAGC CTCCAGCACA GGTGCTTGAA GGCCCTGCGG
2451 TCCGGGCCGG CCATGACCGA GATCGGCGAG CAGCCGTGGG GGCGGGAGTT
     AGGCCCGGCC GGTACTGGCT CTAGCCGCTC GTCGGCACCC CCGCCCTCAA
2501 CGCCCTGCGC GACCCGGCCG GCAACTGCGT GCACTTCGTG GCCGAGGAGC
     GCGGGACGCG CTGGGCCGGC CGTTGACGCA CGTGAAGCAC CGGCTCCTCG
2551 AGGACTGACA CGTGCTACGA GATTTCGATT CCACCGCCGC CTTCTATGAA
     TCCTGACTGT GCACGATGCT CTAAAGCTAA GGTGGCGGCG GAAGATACTT
2601 AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA TGATCCTCCA
     TCCAACCCGA AGCCTTAGCA AAAGGCCCTG CGGCCGACCT ACTAGGAGGT
2651 GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCCAAC TTGTTTATTG
     CGCGCCCCTA GAGTACGACC TCAAGAAGCG GGTGGGGTTG AACAAATAAC
2701 CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT
     GTCGAATATT ACCAATGTTT ATTTCGTTAT CGTAGTGTTT AAAGTGTTTA
2751 AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA
     TTTCGTAAAA AAAGTGACGT AAGATCAACA CCAAACAGGT TTGAGTAGTT
2801 TGTATCTTAT CATGTCTGTA TACCGTCGAC CTCTAGCTAG AGCTTGGCGT
     ACATAGAATA GTACAGACAT ATGGCAGCTG GAGATCGATC TCGAACCGCA
2851 AATCATGGTC ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAATT
     TTAGTACCAG TATCGACAAA GGACACACTT TAACAATAGG CGAGTGTTAA
2901 CCACACAACA TACGAGCCGG AAGCATAAAG TGTAAAGCCT GGGGTGCCTA
     GGTGTGTTGT ATGCTCGGCC TTCGTATTTC ACATTTCGGA CCCCACGGAT
2951 ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTTCC
     TACTCACTCG ATTGAGTGTA ATTAACGCAA CGCGAGTGAC GGGCGAAAGG
3001 AGTCGGGAAA CCTGTCGTGC CAGCTGCATT AATGAATCGG CCAACGCGCG
     TCAGCCCTTT GGACAGCACG GTCGACGTAA TTACTTAGCC GGTTGCGCGC
3051 GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA
     CCCTCTCCGC CAAACGCATA ACCCGCGAGA AGGCGAAGGA GCGAGTGACT
3101 CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA
     GAGCGACGCG AGCCAGCAAG CCGACGCCGC TCGCCATAGT CGAGTGAGTT
3151 AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC
     TCCGCCATTA TGCCAATAGG TGTCTTAGTC CCCTATTGCG TCCTTTCTTG
3201 ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT
     TACACTCGTT TTCCGGTCGT TTTCCGGTCC TTGGCATTTT TCCGGCGCAA
3251 GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC
     CGACCGCAAA AAGGTATCCG AGGCGGGGGG ACTGCTCGTA GTGTTTTTAG
3301 GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG
     CTGCGAGTTC AGTCTCCACC GCTTTGGGCT GTCCTGATAT TTCTATGGTC
3351 GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC
     CGCAAAGGGG GACCTTCGAG GGAGCACGCG AGAGGACAAG GCTGGGACGG
3401 GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT
     CGAATGGCCT ATGGACAGGC GGAAAGAGGG AAGCCCTTCG CACCGCGAAA
3451 CTCAATGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC
     GAGTTACGAG TGCGACATCC ATAGAGTCAA GCCACATCCA GCAAGCGAGG
3501 AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT
     TTCGACCCGA CACACGTGCT TGGGGGGCAA GTCGGGCTGG CGACGCGGAA
3551 ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
     TAGGCCATTG ATAGCAGAAC TCAGGTTGGG CCATTCTGTG CTGAATAGCG
```

FIG.32D

3601 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC
GTGACCGTCG TCGGTGACCA TTGTCCTAAT CGTCTCGCTC CATACATCCG
3651 GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG
CCACGATGTC TCAAGAACTT CACCACCGGA TTGATGCCGA TGTGATCTTC
3701 GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA
CTGTCATAAA CCATAGACGC GAGACGACTT CGGTCAATGG AAGCCTTTTT
3751 GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT
CTCAACCATC GAGAACTAGG CCGTTTGTTT GGTGGCGACC ATCGCCACCA
3801 TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA
AAAAAACAAA CGTTCGTCGT CTAATGCGCG TCTTTTTTTC CTAGAGTTCT
3851 AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT
TCTAGGAAAC TAGAAAAGAT GCCCCAGACT GCGAGTCACC TTGCTTTTGA
3901 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG
GTGCAATTCC CTAAAACCAG TACTCTAATA GTTTTTCCTA GAAGTGGATC
3951 ATCCTTTTAA ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA
TAGGAAAATT TAATTTTTAC TTCAAAATTT AGTTAGATTT CATATATACT
4001 GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG GCACCTATCT
CATTTGAACC AGACTGTCAA TGGTTACGAA TTAGTCACTC CGTGGATAGA
4051 CAGCGATCTG TCTATTTCGT TCATCCATAG TTGCCTGACT CCCCGTCGTG
GTCGCTAGAC AGATAAAGCA AGTAGGTATC AACGGACTGA GGGGCAGCAC
4101 TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT
ATCTATTGAT GCTATGCCCT CCCGAATGGT AGACCGGGGT CACGACGTTA
4151 GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
CTATGGCGCT CTGGGTGCGA GTGGCCGAGG TCTAAATAGT CGTTATTTGG
4201 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC
TCGGTCGGCC TTCCCGGCTC GCGTCTTCAC CAGGACGTTG AAATAGGCGG
4251 TCCATCCAGT CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC
AGGTAGGTCA GATAATTAAC AACGGCCCTT CGATCTCATT CATCAAGCGG
4301 AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC ATCGTGGTGT
TCAATTATCA AACGCGTTGC AACAACGGTA ACGATGTCCG TAGCACCACA
4351 CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA
GTGCGAGCAG CAAACCATAC CGAAGTAAGT CGAGGCCAAG GGTTGCTAGT
4401 AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT
TCCGCTCAAT GTACTAGGGG GTACAACACG TTTTTTCGCC AATCGAGGAA
4451 CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
GCCAGGAGGC TAGCAACAGT CTTCATTCAA CCGGCGTCAC AATAGTGAGT
4501 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA
ACCAATACCG TCGTGACGTA TTAAGAGAAT GACAGTACGG TAGGCATTCT
4551 TGCTTTTCTG TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG
ACGAAAAGAC ACTGACCACT CATGAGTTGG TTCAGTAAGA CTCTTATCAC
4601 TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAATACGG GATAATACCG
ATACGCCGCT GGCTCAACGA GAACGGGCCG CAGTTATGCC CTATTATGGC
4651 CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG
GCGGTGTATC GTCTTGAAAT TTTCACGAGT AGTAACCTTT TGCAAGAAGC
4701 GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA
CCCGCTTTTG AGAGTTCCTA GAATGGCGAC AACTCTAGGT CAAGCTACAT
4751 ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
TGGGTGAGCA CGTGGGTTGA CTAGAAGTCG TAGAAAATGA AAGTGGTCGC

FIG.32E

```
4801 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA
     AAAGACCCAC TCGTTTTTGT CCTTCCGTTT TACGGCGTTT TTTCCCTTAT
4851 AGGGCGACAC GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA
     TCCCGCTGTG CCTTTACAAC TTATGAGTAT GAGAAGGAAA AAGTTATAAT
4901 TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTTGAAT
     AACTTCGTAA ATAGTCCCAA TAACAGAGTA CTCGCCTATG TATAAACTTA
4951 GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA
     CATAAATCTT TTTATTTGTT TATCCCCAAG GCGCGTGTAA AGGGGCTTTT
5001 GTGCCACCTG ACGTC
     CACGGTGGAC TGCAG
```

FIG.32F

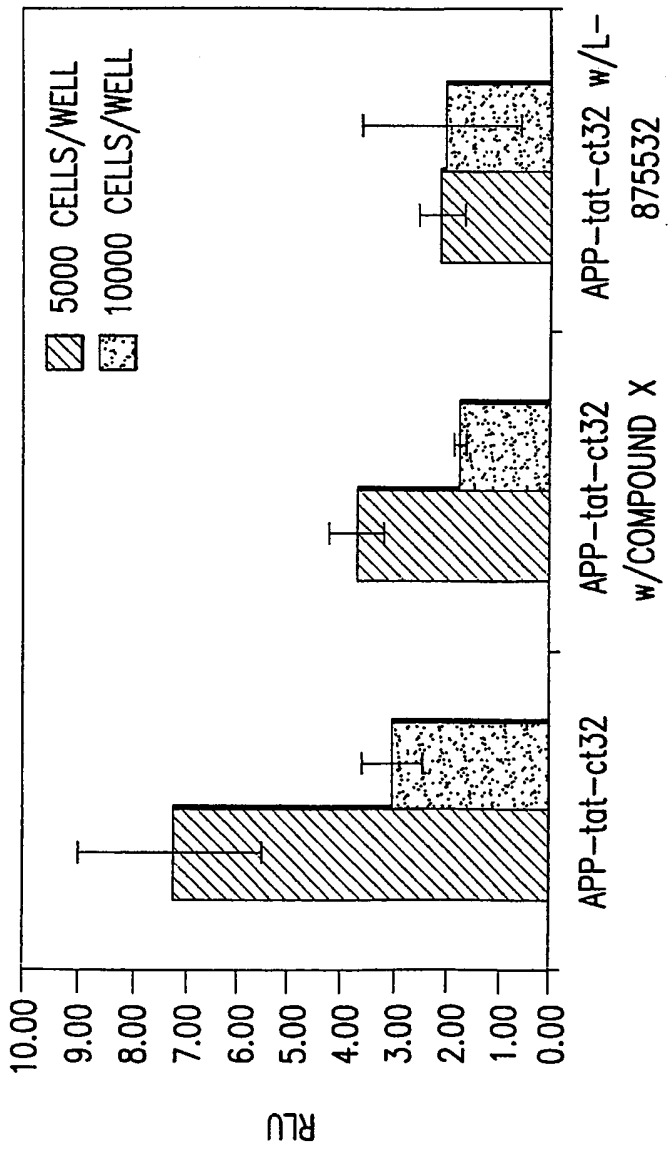

5000 CELLS/WELL
10000 CELLS/WELL
RLU
10.00
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
APP-tat-ct32
APP-tat-ct32 w/COMPOUND X
APP-tat-ct32 w/L-875532

FIG.33

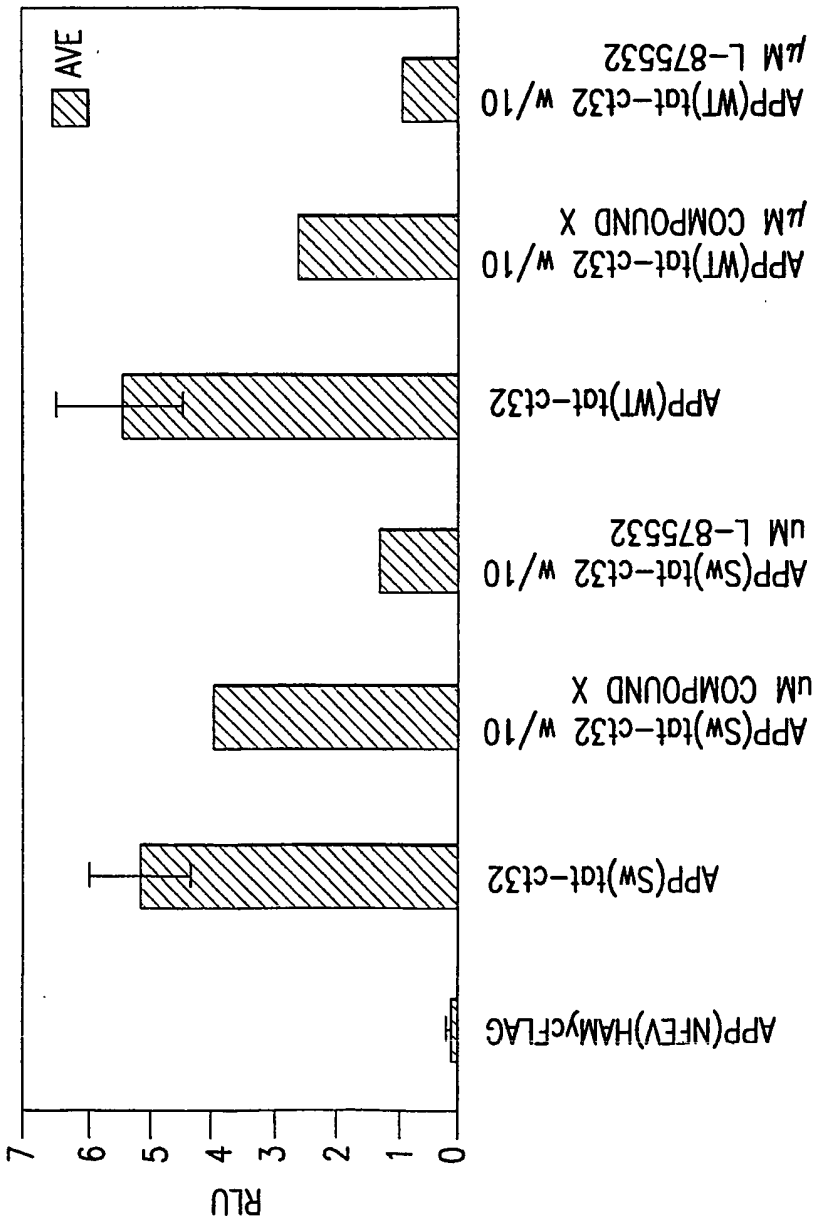

RLU
0
1
2
3
4
5
6
7
AVE
APP(NFEV)HAMycFLAG
APP(Sw)tat-ct32
APP(Sw)tat-ct32 w/10 uM COMPOUND X
APP(Sw)tat-ct32 w/10 uM L-875532
APP(WT)tat-ct32
APP(WT)tat-ct32 w/10 μM COMPOUND X
APP(WT)tat-ct32 w/10 μM L-875532
PLASMID TRANSFECTED

FIG.34

DNA sequence of APP(1-651)NFEV, TATexon1(M1L) APP (664-695)
(SEQ ID NO: 29)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC
1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
```

FIG.35A

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAACTTTGA AGTGGAATTC
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC
2051 CAGGAAGTCA GCCTAAAACT GCTTGTACCA ATTGCTATTG TAAAAAGTGT
2101 TGCTTTCATT GCCAAGTTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG.35B

(SEQ ID NO: 30)

Amino acid sequence of APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

1

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefslddlqpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

2 3 4 5

nfevefrhdsgyevhhqklvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

6

~~adi~~lepvdprlepwkhpgsqpktactncyckkccfhcqvcfmtkalqisygrkkrrqrrrahqnsqthqaslskq

7 8

~~risstvaaa~~daavtpeerhlskmqqnqyenptykffeqmqn

FIG.36

DNA sequence of APP(1-651)NFEV, GAL4VP16(delMet) APP (664-695)
(SEQ ID NO: 31)

```
   1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
  51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
 101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAATGGG
 151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
 201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
 251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
 301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
 351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCT GACAAGTGCA
 401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
 451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
 501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
 551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
 601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
 651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
 701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
 751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
 801 GAGAACCACC AGCATTGCCA CCACCACCAC CACCACCACA GAGTCTGTGG
 851 AAGAGGTGGT TCGAGTTCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
 901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
 951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCTCCT CGGCCTCGTC ACGTGTTCAA
```

FIG.37A

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACA GAAGGACAGA CAGCACACCC

1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATTC

1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC

1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC

1551 TCTCATGCCA TCTTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG

1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTTGGG

1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC

1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA

1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAACTTTGA AGTGGAATTC

1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTTGC

1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG

1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG

1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTC

2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTTGCC

2051 GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT

2101 CTGAAGAACA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC

2151 GCTGACTAGG GCACATCTGA CAGAAGTGGA ATCAAGGCTA GAAAGACTGG

2201 AACAGCTATT TCTACTGATT TTTCCTCGAG AAGACCTTGA CATGATTTTG

2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT

2301 ACAAGATAAT GTGAATAAAG ATGCCGTCAC AGATAGATTG GCTTCAGTGG

2351 AGACTGATAT GCCTCTAACA TTGAGACAGC ATAGAATAAG TGCGACATCA

2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG

2451 AATTCCCGGG GATCTGGCCC CCCCGACCGA TGTCAGCCTG GGGGACGAGC

2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC

2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.37B

2601 TACCCCCCAC GACTCCGCCC CCTACGGCGC TCTGGATATG GCCGACTTCG

2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG

2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG

2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT

2801 TCTTTGAGCA GATGCAGAAC TAG

FIG.37C

(SEQ ID NO: 32)

Amino acid sequence of APP(1-651)NFEV, GAL4-VP16(delM1) APP (664-695)

mlpglallllaawtaralevptdgnagllaepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq

evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh

whtvaketcsekstnlhdygmllpcgidkfrgvefvccplaeesdnvdsadaeeddsdvwwggadtdyadgs

edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttttesveevvrvpttaastpd

avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqaknlpkadkkaviqhfqekvesleqe

aanerqqlvethmarveamlndrrrlalenyitalqavpprprhvfnmlkkyvraeqkdrqhtlkhfehvrmvd

pkkaaqirsqvmthlrviyermnqslsllynvpavaeeiqdevdellqkeqnysddvlanmiseprisygndal

mpsltetkttvellpvngefslddlqpwhsfgadsvpantenevepvdarpaadrglttrpgsgltnikteeisev

nfevefrhdsgyevhhqklvffaedvgsnkgaiiglmvggvviatvivitlvmlkkk~~klgtelgstspvwwns~~

~~adi~~kllssieqacdicrlkklkcskekpkcakclknnwecryspktkrspltrahltevesrlerleqlfllifpredld

milkmdslqdikalltqlfvqdnvnkdavtdrlasvetdmpltlrqhrisatssseessnkqqrqltvsqipqdlapp

tdvslqdelhldqedvamahadaldqfdldmlqdqdspqpqftphdsapyqaldmadfefeqmftdalqidey

~~ggdiqhsgaaa~~daavtpeerhlskmqqnqyenptykffeqmqn

FIG.38

SEQUENCE LISTING

<110> Merck & Co., Inc.
Espeseth, Amy S.
Ferrer, Marc
Flores, Osvaldo A.
Hazuda, Daria J.
Inglese, James
Miller, Michael D.
Register, Bruce
Shi, Xiao-Ping
Simon, Adam J.
Zuck, Paul D.

<120> ASSAYS TO MONITOR AMYLOID PRECURSOR PROTEIN PROCESSING

<130> 21040-PCT

<150> 60/360,274
<151> 2002-02-27

<160> 64

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 2352
<212> DNA
<213> fusion protein - human

<400> 1
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740

tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctaga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caagtattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc ctggagccag tagatcctag actagagccc 2040
tggaagcatc caggaagtca gcctaaaact gcttgtacca attgctattg taaaaagtgt 2100
tgctttcatt gccaagtttg tttcatgaca aaagccttag gcatctccta tggcaggaag 2160
aagcggagac agcgacgaag agctcatcag aacagtcaga ctcatcaagc ttctctatca 2220
aagcagagga tatccagcac agtggcggcc gcagacgccg ctgtcacccc agaggagcgc 2280
cacctgtcca agatgcagca gaacggctac gaaaatccaa cctacaagtt ctttgagcag 2340
atgcagaact ag 2352

<210> 2
<211> 783
<212> PRT
<213> fusion protein - human

<400> 2

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
```

```
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
            660                 665                 670
Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
        675                 680                 685
Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
    690                 695                 700
Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
705                 710                 715                 720
Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr His Gln
                725                 730                 735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Ala Asp
            740                 745                 750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
        755                 760                 765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    770                 775                 780
```

<210> 3
<211> 2352
<212> DNA
<213> fusion protein - human

<400> 3

atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caagtattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc ctggagccag tagatcctag actagagccc 2040
tggaagcatc caggaagtca gcctaaaact gcttgtacca attgctattg taaaaagtgt 2100
tgctttcatt gccaagtttg tttcatgaca aaagccttag gcatctccta tggcaggaag 2160
aagcggagac agcgacgaag agctcatcag aacagtcaga ctcatcaagc ttctctatca 2220
aagcagagga tatccagcac agtggcggcc gcagacgccg ctgtcacccc agaggagcgc 2280
cacctgtcca agatgcagca gaacggctac gaaaatccaa cctacaagtt ctttgagcag 2340
atgcagaact ag 2352

<210> 4
<211> 783
<212> PRT
<213> fusion protein - human

<400> 4

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
```

```
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
```

```
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
            660                 665                 670
Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
        675                 680                 685
Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
    690                 695                 700
Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
705                 710                 715                 720
Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr His Gln
                725                 730                 735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Ala Asp
            740                 745                 750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
        755                 760                 765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    770                 775                 780
```

<210> 5
<211> 2823
<212> DNA
<213> fusion protein - human

<400> 5

```
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctaga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caagtattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
```

agtccagtgt ggtggaattc tgcagatatc aagctactgt cttctatcga acaagcatgc 2040
gatatttgcc gacttaaaaa gctcaagtgc tccaaagaaa aaccgaagtg cgccaagtgt 2100
ctgaagaaca actgggagtg tcgctactct cccaaaacca aaaggtctcc gctgactagg 2160
gcacatctga cagaagtgga atcaaggcta gaaagactgg aacagctatt tctactgatt 2220
tttcctcgag aagaccttga catgattttg aaaatggatt ctttacagga tataaaagca 2280
ttgttaacag gattatttgt acaagataat gtgaataaag atgccgtcac agatagattg 2340
gcttcagtgg agactgatat gcctctaaca ttgagacagc atagaataag tgcgacatca 2400
tcatcggaag agagtagtaa caaaggtcaa agacagttga ctgtatcggg aattcccggg 2460
gatctggccc ccccgaccga tgtcagcctg ggggacgagc tccacttaga cggcgaggac 2520
gtggcgatgg cgcatgccga cgcgctagac gatttcgatc tggacatgtt gggggacggg 2580
gattccccgg gtccgggatt tacccccac gactccgccc cctacggcgc tctggatatg 2640
gccgacttcg agtttgagca gatgtttacc gatgcccttg gaattgacga gtacggtggg 2700
gatatccagc acagtggcgg ccgcgacgcc gctgtcaccc cagaggagcg ccacctgtcc 2760
aagatgcagc agaacggcta cgaaaatcca acctacaagt tctttgagca gatgcagaac 2820
tag 2823

<210> 6
<211> 941
<212> PRT
<213> fusion protein - human

<400> 6

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
```

```
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
            660                 665                 670
Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
        675                 680                 685
Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
    690                 695                 700
Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
705                 710                 715                 720
Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
                725                 730                 735
Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
            740                 745                 750
Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
        755                 760                 765
Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
    770                 775                 780
Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
785                 790                 795                 800
Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
                805                 810                 815
```

```
Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
            820                 825                 830
Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
        835                 840                 845
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
    850                 855                 860
Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
865                 870                 875                 880
Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
                885                 890                 895
Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala
            900                 905                 910
Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
        915                 920                 925
Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    930                 935                 940
```

<210> 7
<211> 2823
<212> DNA
<213> fusion protein - human

<400> 7

```
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caagtattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc aagctactgt cttctatcga acaagcatgc 2040
gatatttgcc gacttaaaaa gctcaagtgc tccaaagaaa aaccgaagtg cgccaagtgt 2100
ctgaagaaca actgggagtg tcgctactct cccaaaacca aaaggtctcc gctgactagg 2160
gcacatctga cagaagtgga atcaaggcta gaaagactgg aacagctatt tctactgatt 2220
tttcctcgag aagaccttga catgattttg aaaatggatt ctttacagga tataaaagca 2280
```

```
ttgttaacag gattatttgt acaagataat gtgaataaag atgccgtcac agatagattg 2340
gcttcagtgg agactgatat gcctctaaca ttgagacagc atagaataag tgcgacatca 2400
tcatcggaag agagtagtaa caaaggtcaa agacagttga ctgtatcggg aattcccggg 2460
gatctggccc ccccgaccga tgtcagcctg ggggacgagc tccacttaga cggcgaggac 2520
gtggcgatgg cgcatgccga cgcgctagac gatttcgatc tggacatgtt gggggacggg 2580
gattccccgg gtccgggatt tacccccccac gactccgccc cctacggcgc tctggatatg 2640
gccgacttcg agtttgagca gatgtttacc gatgcccttg gaattgacga gtacggtggg 2700
gatatccagc acagtggcgg ccgcgacgcc gctgtcaccc cagaggagcg ccacctgtcc 2760
aagatgcagc agaacggcta cgaaaatcca acctacaagt tctttgagca gatgcagaac 2820
tag                                                               2823

<210> 8
<211> 941
<212> PRT
<213> fusion protein - human

<400> 8
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
```

```
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
            660                 665                 670
Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
        675                 680                 685
Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
    690                 695                 700
Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
705                 710                 715                 720
Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
                725                 730                 735
Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
            740                 745                 750
Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
        755                 760                 765
Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
    770                 775                 780
Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
785                 790                 795                 800
Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
                805                 810                 815
Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
            820                 825                 830
Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
        835                 840                 845
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
```

```
    850                 855                 860
Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
865                 870                 875                 880
Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
                885                 890                 895
Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala
            900                 905                 910
Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
        915                 920                 925
Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    930                 935                 940
```

<210> 9
<211> 2352
<212> DNA
<213> fusion protein - human

<400> 9

```
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctaga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc ctggagccag tagatcctag actagagccc 2040
tggaagcatc caggaagtca gcctaaaact gcttgtacca attgctattg taaaaagtgt 2100
tgctttcatt gccaagtttg tttcatgaca aaagccttag gcatctccta tggcaggaag 2160
aagcggagac agcgacgaag agctcatcag aacagtcaga ctcatcaagc ttctctatca 2220
aagcagagga tatccagcac agtggcggcc gcagacgccg ctgtcacccc agaggagcgc 2280
cacctgtcca agatgcagca gaacggctac gaaaatccaa cctacaagtt ctttgagcag 2340
atgcagaact ag                                                    2352
```

<210> 10
<211> 783

<212> PRT
<213> fusion protein - human

<400> 10

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
```

```
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
            660                 665                 670
Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
        675                 680                 685
Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
    690                 695                 700
Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
705                 710                 715                 720
Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr His Gln
                725                 730                 735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Ala Asp
            740                 745                 750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
        755                 760                 765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    770                 775                 780
```

<210> 11
<211> 2352
<212> DNA
<213> fusion protein - human

<400> 11
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840

gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc ctggagccag tagatcctag actagagccc 2040
tggaagcatc caggaagtca gcctaaaact gcttgtacca attgctattg taaaaagtgt 2100
tgctttcatt gccaagtttg tttcatgaca aaagccttag gcatctccta tggcaggaag 2160
aagcggagac agcgacgaag agctcatcag aacagtcaga ctcatcaagc ttctctatca 2220
aagcagagga tatccagcac agtggcggcc gcagacgccg ctgtcacccc agaggagcgc 2280
cacctgtcca agatgcagca gaacggctac gaaaatccaa cctacaagtt ctttgagcag 2340
atgcagaact ag 2352

<210> 12
<211> 783
<212> PRT
<213> fusion protein - human

<400> 12

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
```

```
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
            660                 665                 670
Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
        675                 680                 685
Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
    690                 695                 700
Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
705                 710                 715                 720
Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr His Gln
```

```
                725                 730                 735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Asp
            740                 745                 750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
        755                 760                 765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    770                 775                 780
```

<210> 13
<211> 2823
<212> DNA
<213> fusion protein - human

<400> 13

```
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctaga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc aagctactgt cttctatcga acaagcatgc 2040
gatatttgcc gacttaaaaa gctcaagtgc tccaaagaaa aaccgaagtg cgccaagtgt 2100
ctgaagaaca actgggagtg tcgctactct cccaaaacca aaaggtctcc gctgactagg 2160
gcacatctga cagaagtgga atcaaggcta gaaagactgg aacagctatt tctactgatt 2220
tttcctcgag aagaccttga catgattttg aaaatggatt ctttacagga tataaaagca 2280
ttgttaacag gattatttgt acaagataat gtgaataaag atgccgtcac agatagattg 2340
gcttcagtgg agactgatat gcctctaaca ttgagacagc atagaataag tgcgacatca 2400
tcatcggaag agagtagtaa caaaggtcaa agacagttga ctgtatcggg aattcccggg 2460
gatctggccc ccccgaccga tgtcagcctg ggggacgagc tccacttaga cggcgaggac 2520
gtggcgatgg cgcatgccga cgcgctagac gatttcgatc tggacatgtt gggggacggg 2580
gattccccgg gtccgggatt tacccccac gactccgccc cctacggcgc tctggatatg 2640
gccgacttcg agtttgagca gatgtttacc gatgcccttg gaattgacga gtacggtggg 2700
gatatccagc acagtggcgg ccgcgacgcc gctgtcaccc cagaggagcg ccacctgtcc 2760
aagatgcagc agaacggcta cgaaaatcca acctacaagt tctttgagca gatgcagaac 2820
```

tag 2823

<210> 14
<211> 941
<212> PRT
<213> fusion protein - human

<400> 14

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
```

```
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
            660                 665                 670
Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
        675                 680                 685
Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
    690                 695                 700
Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
705                 710                 715                 720
Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
                725                 730                 735
Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
            740                 745                 750
Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
        755                 760                 765
Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
    770                 775                 780
Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
785                 790                 795                 800
Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
                805                 810                 815
Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
            820                 825                 830
Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
        835                 840                 845
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
    850                 855                 860
Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
865                 870                 875                 880
Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
                885                 890                 895
Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala
            900                 905                 910
Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
        915                 920                 925
```

Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
930 935 940

<210> 15
<211> 2823
<212> DNA
<213> fusion protein - human

<400> 15
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc aagctactgt cttctatcga acaagcatgc 2040
gatatttgcc gacttaaaaa gctcaagtgc tccaaagaaa aaccgaagtg cgccaagtgt 2100
ctgaagaaca actgggagtg tcgctactct cccaaaacca aaaggtctcc gctgactagg 2160
gcacatctga cagaagtgga atcaaggcta gaaagactgg aacagctatt tctactgatt 2220
tttcctcgag aagaccttga catgattttg aaaatggatt ctttacagga tataaaagca 2280
ttgttaacag gattatttgt acaagataat gtgaataaag atgccgtcac agatagattg 2340
gcttcagtgg agactgatat gcctctaaca ttgagacagc atagaataag tgcgacatca 2400
tcatcggaag agagtagtaa caaaggtcaa agacagttga ctgtatcggg aattcccggg 2460
gatctggccc ccccgaccga tgtcagcctg ggggacgagc tccacttaga cggcgaggac 2520
gtggcgatgg cgcatgccga cgcgctagac gatttcgatc tggacatgtt gggggacggg 2580
gattccccgg gtccgggatt tacccccac gactccgccc cctacggcgc tctggatatg 2640
gccgacttcg agtttgagca gatgtttacc gatgcccttg gaattgacga gtacggtggg 2700
gatatccagc acagtggcgg ccgcgacgcc gctgtcaccc cagaggagcg ccacctgtcc 2760
aagatgcagc agaacggcta cgaaaatcca acctacaagt tctttgagca gatgcagaac 2820
tag 2823

<210> 16
<211> 941
<212> PRT

<213> fusion protein - human

<400> 16

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
```

```
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
            660                 665                 670
Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
        675                 680                 685
Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
    690                 695                 700
Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
705                 710                 715                 720
Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
                725                 730                 735
Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
            740                 745                 750
Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
        755                 760                 765
Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
    770                 775                 780
Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
785                 790                 795                 800
Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
                805                 810                 815
Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
            820                 825                 830
Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
        835                 840                 845
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
    850                 855                 860
Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
865                 870                 875                 880
Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
                885                 890                 895
Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala
            900                 905                 910
Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
        915                 920                 925
Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    930                 935                 940
```

<210> 17

<211> 3354
<212> DNA
<213> human

<400> 17
agtttcctcg gcagcggtag gcgagagcac gcggaggagc gtgcgcgggg gccccgggag 60
acggcggcgg tggcggcgcg ggcagagcaa ggacgcggcg gatcccactc gcacagcagc 120
gcactcggtg ccccgcgcag ggtcgcgatg ctgcccggtt tggcactgct cctgctggcc 180
gcctggacgg ctcgggcgct ggaggtaccc actgatggta atgctggcct gctggctgaa 240
ccccagattg ccatgttctg tggcagactg aacatgcaca tgaatgtcca gaatgggaag 300
tgggattcag atccatcagg gaccaaaacc tgcattgata ccaaggaagg catcctgcag 360
tattgccaag aagtctaccc tgaactgcag atcaccaatg tggtagaagc caaccaacca 420
gtgaccatcc agaactggtg caagcggggc cgcaagcagt gcaagaccca tccccacttt 480
gtgattccct accgctgctt agttggtgag tttgtaagtg atgcccttct cgttcctgac 540
aagtgcaaat tcttacacca ggagaggatg gatgtttgcg aaactcatct tcactggcac 600
accgtcgcca aagagacatg cagtgagaag agtaccaact tgcatgacta cggcatgttg 660
ctgccctgcg gaattgacaa gttccgaggg gtagagtttg tgtgttgccc actggctgaa 720
gaaagtgaca atgtggattc tgctgatgcg gaggaggatg actcggatgt ctggtggggc 780
ggagcagaca cagactatgc agatgggagt gaagacaaag tagtagaagt agcagaggag 840
gaagaagtgg ctgaggtgga agaagaagaa gccgatgatg acgaggacga tgaggatggt 900
gatgaggtag aggaagaggc tgaggaaccc tacgaagaag ccacagagag aaccaccagc 960
attgccacca ccaccaccac caccacagag tctgtggaag aggtggttcg agttcctaca 1020
acagcagcca gtacccctga tgccgttgac aagtatctcg agacacctgg ggatgagaat 1080
gaacatgccc atttccagaa agccaaagag aggcttgagg ccaagcaccg agagagaatg 1140
tcccaggtca tgagagaatg ggaagaggca gaacgtcaag caaagaactt gcctaaagct 1200
gataagaagg cagttatcca gcatttccag gagaaagtgg aatctttgga acaggaagca 1260
gccaacgaga gacagcagct ggtggagaca cacatggcca gagtggaagc catgctcaat 1320
gaccgccgcc gcctggccct ggagaactac atcaccgctc tgcaggctgt tcctcctcgg 1380
cctcgtcacg tgttcaatat gctaaagaag tatgtccgcg cagaacagaa ggacagacag 1440
cacaccctaa agcatttcga gcatgtgcgc atggtggatc ccaagaaagc cgctcagatc 1500
cggtcccagg ttatgacaca cctccgtgtg atttatgagc gcatgaatca gtctctctcc 1560
ctgctctaca acgtgcctgc agtggccgag gagattcagg atgaagttga tgagctgctt 1620
cagaaagagc aaaactattc agatgacgtc ttggccaaca tgattagtga accaaggatc 1680
agttacggaa acgatgctct catgccatct ttgaccgaaa cgaaaaccac cgtggagctc 1740
cttcccgtga atggagagtt cagcctggac gatctccagc cgtggcattc ttttggggct 1800
gactctgtgc cagccaacac agaaaacgaa gttgagcctg ttgatgcccg ccctgctgcc 1860
gaccgaggac tgaccactcg accaggttct gggttgacaa atatcaagac ggaggagatc 1920
tctgaagtga agatggatgc agaattccga catgactcag gatatgaagt tcatcatcaa 1980
aaattggtgt tctttgcaga agatgtgggt tcaaacaaag gtgcaatcat tggactcatg 2040
gtgggcggtg ttgtcatagc gacagtgatc gtcatcacct tggtgatgct gaagaagaaa 2100
cagtacacat ccattcatca tggtgtggtg gaggttgacg ccgctgtcac cccagaggag 2160
cgccacctgt ccaagatgca gcagaacggc tacgaaaatc caacctacaa gttctttgag 2220
cagatgcaga actagacccc cgccacagca gcctctgaag ttggacagca aaaccattgc 2280
ttcactaccc atcggtgtcc atttatagaa taatgtggga agaaacaaac ccgttttatg 2340
atttactcat tatcgccttt tgacagctgt gctgtaacac aagtagatgc ctgaacttga 2400
attaatccac acatcagtaa tgtattctat ctctctttac attttggtct ctatactaca 2460
ttattaatgg gttttgtgta ctgtaaagaa tttagctgta tcaaactagt gcatgaatag 2520
attctctcct gattatttat cacatagccc cttagccagt tgtatattat tcttgtggtt 2580
tgtgacccaa ttaagtccta ctttacatat gctttaagaa tcgatggggg atgcttcatg 2640
tgaacgtggg agttcagctg cttctcttgc ctaagtattc ctttcctgat cactatgcat 2700
tttaaagtta aacatttta agtatttcag atgctttaga gagatttttt ttccatgact 2760
gcattttact gtacagattg ctgcttctgc tatatttgtg atataggaat taagaggata 2820
cacacgtttg tttcttcgtg cctgttttat gtgcacacat taggcattga gacttcaagc 2880
ttttcttttt ttgtccacgt atctttgggt ctttgataaa gaaaagaatc cctgttcatt 2940
gtaagcactt ttacggggcg ggtggggagg ggtgctctgc tggtcttcaa ttaccaagaa 3000
ttctccaaaa caattttctg caggatgatt gtacagaatc attgcttatg acatgatcgc 3060
tttctacact gtattacata aataaattaa ataaaataac cccgggcaag acttttcttt 3120
gaaggatgac tacagacatt aaataatcga agtaattttg ggtggggaga agaggcagat 3180
tcaattttct ttaaccagtc tgaagtttca tttatgatac aaaagaagat gaaaatggaa 3240
gtggcaatat aaggggatga ggaaggcatg cctggacaaa cccttctttt aagatgtgtc 3300
ttcaatttgt ataaaatggt gttttcatgt aaataaatac attcttggag gagc 3354

<210> 18
<211> 695
<212> PRT
<213> human

<400> 18

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
```

```
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
                645                 650                 655
His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
            660                 665                 670
His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
        675                 680                 685
Phe Phe Glu Gln Met Gln Asn
    690                 695
```

<210> 19
<211> 3972
<212> DNA
<213> human

<400> 19

```
atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60
ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120
ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180
ctcgtgacca ccttcaccta cggcgtgcag tgcttcgccc gctaccccga ccacatgaag 240
cagcacgact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc 300
ttcaaggacg acggcaacta caagacccgc gccgaggtga agttcgaggg cgacaccctg 360
gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420
aagctggagt acaactacaa cagccacaag gtctatatca ccgccgacaa gcagaagaac 480
ggcatcaagg tgaacttcaa gacccgccac aacatcgagg acggcagcgt gcagctcgcc 540
gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600
tacctgagca cccagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatggtc 660
ctgctggagt tcgtgaccgc cgccgggatc actctcggca tggacgagct gtacaagtaa 720
ctcgagtcta gagggcccgt ttaaacccgc tgatcagcct cgactgtgcc ttctagttgc 780
cagccatctg ttgtttgccc ctcccccgtg ccttccttga ccctggaagg tgccactccc 840
actgtccttt cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct 900
attctggggg gtggggtggg gcaggacagc aagggggagg attgggaaga caatagcagg 960
catgctgggg atgcggtggg ctctatggct tctgaggcgg aaagaaccag catgtgagca 1020
aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg 1080
ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 1140
acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 1200
ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 1260
tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 1320
```

tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt 1380
gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 1440
agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc 1500
tacactagaa gaacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa 1560
agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttgtt 1620
tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 1680
acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgacatta 1740
acctataaaa ataggcgtat cacgaggccc tttcgtctcg cgcgtttcgg tgatgacggt 1800
gaaaacctct gacacatgca gctcccggag acggtcacag cttgtctgta agcggatgcc 1860
gggagcagac aagcccgtca gggcgcgtca gcgggtgttg gcgggtgtcg gggctggctt 1920
aactatgcgg catcagagca gattgtactg agagtgcacc atatgcggtg tgaaataccg 1980
cacagatgcg taaggagaaa ataccgcatc aggacgcgcc ctgtagcggc gcattaagcg 2040
cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg 2100
ctcctttcgc tttcttccct tcctttctcg ccacgttcgc cggctttccc cgtcaagctc 2160
taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa 2220
aacttgatta gggtgatggt tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc 2280
ctttgacgtt ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac 2340
tcaaccctat ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt 2400
ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc 2460
ttacaatttc cattcgccat tcaggctgaa ctagatctag agtccgttac ataacttacg 2520
gtaaatggcc cgcctggctg accgcccaac gacccccgcc cattgacgtc aataatgacg 2580
tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta 2640
cggtaaactg cccacttggc agtacatcaa gtgtatcata tgccaagtac gcccccctatt 2700
gacgtcaatg acggtaaatg gcccgcctgg cattatgccc agtacatgac cttatgggac 2760
tttcctactt ggcagtacat ctacgtatta gtcatcgcta ttaccatggt gatgcggttt 2820
tggcagtaca tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac 2880
cccattgacg tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt 2940
cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc gtgtacggtg ggaggtctat 3000
ataagcagag ctcgtttagt gaaccgtcag atcgcctgga gacgccatcc acgctgtttt 3060
gacctccata gaagacaccg ggaccgatcc agcctccgcg gccgggaacg gtgcattgga 3120
acggaccgtg ttgacaatta atcatcggca tagtatatcg gcatagtata atacgacaag 3180
gtgaggaact aaaccatggc caagttgacc agtgccgttc cggtgctcac cgcgcgcgac 3240
gtcgccggag cggtcgagtt ctggaccgac cggctcgggt tctcccggga cttcgtggag 3300
gacgacttcg ccggtgtggt ccgggacgac gtgaccctgt tcatcagcgc ggtccaggac 3360
caggtggtgc cggacaacac cctggcctgg gtgtgggtgc gcggcctgga cgagctgtac 3420
gccgagtggt cggaggtcgt gtccacgaac ttccgggacg cctccgggcc ggccatgacc 3480
gagatcggcg agcagccgtg ggggcgggag ttcgccctgc gcgacccggc cggcaactgc 3540
gtgcacttcg tggccgagga gcaggactga cactcgacct cgaaacttgt ttattgcagc 3600
ttataatggt tacaaataaa gcaatagcat cacaaatttc acaaataaag catttttttc 3660
actgcattct agttgtggtt tgtccaaact catcaatgta tcttatcatg tctggatcga 3720
tacttcaaga actgctgaca tcgagcttgc tacaagggac tttccgctgg ggactttcca 3780
gggaggcgtg gcctgggcgg gactggggag tggcgagccc tcagatcctg catataagca 3840
gctgcttttt gcctgtactg ggtctctctg gttagaccag atctgagcct gggagctctc 3900
tggctaacta gggaacccac tgcttaagcc tcaataaagc ttggtaccga gctcggatcc 3960
gaattcgcca cc 3972

<210> 20
<211> 3972
<212> DNA
<213> human

<400> 20
taccactcgt tcccgctcct cgacaagtgg ccccaccacg ggtaggacca gctcgacctg 60
ccgctgcatt tgccggtgtt caagtcgcac aggccgctcc cgctcccgct acggtggatg 120
ccgttcgact gggacttcaa gtagacgtgg tggccgttcg acgggcacgg gaccgggtgg 180
gagcactggt ggaagtggat gccgcacgtc acgaagcggg cgatggggct ggtgtacttc 240
gtcgtgctga agaagttcag gcggtacggg cttccgatgc aggtcctcgc gtggtagaag 300
aagttcctgc tgccgttgat gttctgggcg cggctccact tcaagctccc gctgtgggac 360
cacttggcgt agctcgactt cccgtagctg aagttcctcc tgccgttgta ggaccccgtg 420
ttcgacctca tgttgatgtt gtcggtgttc cagatatagt ggcggctgtt cgtcttcttg 480
ccgtagttcc acttgaagtt ctgggcggtg ttgtagctcc tgccgtcgca cgtcgagcgg 540
ctggtgatgg tcgtcttgtg ggggtagccg ctgccggggc acgacgacgg gctgttggtg 600

atggactcgt gggtcaggcg ggactcgttt ctggggttgc tcttcgcgct agtgtaccag 660
gacgacctca agcactggcg gcggccctag tgagagccgt acctgctcga catgttcatt 720
gagctcagat ctcccgggca aatttgggcg actagtcgga gctgacacgg aagatcaacg 780
gtcggtagac aacaaacggg gaggggggcac ggaaggaact gggaccttcc acggtgaggg 840
tgacaggaaa ggattatttt actcctttaa cgtagcgtaa cagactcatc cacagtaaga 900
taagaccccc caccccaccc cgtcctgtcg ttccccctcc taacccttct gttatcgtcc 960
gtacgacccc tacgccaccc gagataccga agactccgcc tttcttggtc gtacactcgt 1020
tttccggtcg ttttccggtc cttggcattt ttccggcgca acgaccgcaa aaaggtatcc 1080
gaggcggggg gactgctcgt agtgttttta gctgcgagtt cagtctccac cgctttgggc 1140
tgtcctgata tttctatggt ccgcaaaggg ggaccttcga gggagcacgc gagaggacaa 1200
ggctgggacg gcgaatggcc tatggacagg cggaaagagg gaagcccttc gcaccgcgaa 1260
agagtatcga gtgcgacatc catagagtca agccacatcc agcaagcgag gttcgacccg 1320
acacacgtgc ttgggggggca agtcgggctg gcgacgcgga ataggccatt gatagcagaa 1380
ctcaggttgg gccattctgt gctgaatagc ggtgaccgtc gtcggtgacc attgtcctaa 1440
tcgtctcgct ccatacatcc gccacgatgt ctcaagaact tcaccaccgg attgatgccg 1500
atgtgatctt cttgtcataa accatagacg cgagacgact tcggtcaatg gaagcctttt 1560
tctcaaccat cgagaactag gccgtttgtt tggtggcgac catcgccacc aaaaaaacaa 1620
acgttcgtcg tctaatgcgc gtcttttttt cctagagttc ttctaggaaa ctagaaaaga 1680
tgccccagac tgcgagtcac cttgcttttg agtgcaattc cctaaaacca gtactgtaat 1740
tggatatttt tatccgcata gtgctccggg aaagcagagc gcgcaaagcc actactgcca 1800
cttttggaga ctgtgtacgt cgagggcctc tgccagtgtc gaacagacat tcgcctacgg 1860
ccctcgtctg ttcgggcagt cccgcgcagt cgcccacaac cgcccacagc cccgaccgaa 1920
ttgatacgcc gtagtctcgt ctaacatgac tctcacgtgg tatacgccac actttatggc 1980
gtgtctacgc attcctcttt tatggcgtag tcctgcgcgg gacatcgccg cgtaattcgc 2040
gccgcccaca ccaccaatgc gcgtcgcact ggcgatgtga acggtcgcgg gatcgcgggc 2100
gaggaaagcg aaagaaggga aggaaagagc ggtgcaagcg gccgaaaggg gcagttcgag 2160
atttagcccc cgagggaaat cccaaggcta aatcacgaaa tgccgtggag ctggggtttt 2220
ttgaactaat cccactacca agtgcatcac ccggtagcgg gactatctgc caaaaagcgg 2280
gaaactgcaa cctcaggtgc aagaaattat cacctgagaa caaggtttga ccttgttgtg 2340
agttgggata gagccagata agaaaactaa atattcccta aaacggctaa agccggataa 2400
ccaatttttt actcgactaa attgttttta aattgcgctt aaaattgttt tataattgcg 2460
aatgttaaag gtaagcggta agtccgactt gatctagatc tcaggcaatg tattgaatgc 2520
catttaccgg gcggaccgac tggcgggttg ctgggggcgg gtaactgcag ttattactgc 2580
atacaagggt atcattgcgg ttatccctga aaggtaactg cagttaccca cctcataaat 2640
gccatttgac gggtgaaccg tcatgtagtt cacatagtat acggttcatg cgggggataa 2700
ctgcagttac tgccatttac cgggcggacc gtaatacggg tcatgtactg gaataccctg 2760
aaaggatgaa ccgtcatgta gatgcataat cagtagcgat aatggtacca ctacgccaaa 2820
accgtcatgt agttacccgc acctatcgcc aaactgagtg cccctaaagg ttcagaggtg 2880
gggtaactgc agttaccctc aaacaaaacc gtggttttag ttgccctgaa aggttttaca 2940
gcattgttga ggcggggtaa ctgcgtttac ccgccatccg cacatgccac cctccagata 3000
tattcgtctc gagcaaatca cttggcagtc tagcggacct ctgcggtagg tgcgacaaaa 3060
ctggaggtat cttctgtggc cctggctagg tcggaggcgc cggcccttgc cacgtaacct 3120
tgcctggcac aactgttaat tagtagccgt atcatatagc cgtatcatat tatgctgttc 3180
cactccttga tttggtaccg gttcaactgg tcacggcaag gccacgagtg gcgcgcgctg 3240
cagcggcctc gccagctcaa gacctggctg gccgagccca agagggccct gaagcacctc 3300
ctgctgaagc ggccacacca ggccctgctg cactgggaca agtagtcgcg ccaggtcctg 3360
gtccaccacg gcctgttgtg ggaccggacc cacacccacg cgccggacct gctcgacatg 3420
cggctcacca gcctccagca caggtgcttg aaggccctgc ggaggcccgg ccggtactgg 3480
ctctagccgc tcgtcggcac ccccgccctc aagcgggacg cgctgggccg gccgttgacg 3540
cacgtgaagc accggctcct cgtcctgact gtgagctgga gctttgaaca aataacgtcg 3600
aatattacca atgtttattt cgttatcgta gtgtttaaag tgtttatttc gtaaaaaaag 3660
tgacgtaaga tcaacaccaa acaggtttga gtagttacat agaatagtac agacctagct 3720
atgaagttct tgacgactgt agctcgaacg atgttccctg aaaggcgacc cctgaaaggt 3780
ccctccgcac cggacccgcc ctgacccctc accgctcggg agtctaggac gtatattcgt 3840
cgacgaaaaa cggacatgac ccagagagac caatctggtc tagactcgga ccctcgagag 3900
accgattgat cccttgggtg acgaattcgg agttatttcg aaccatggct cgagcctagg 3960
cttaagcggt gg 3972

<210> 21
<211> 7153
<212> DNA
<213> human

<400> 21
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc 900
gtttaaactt aagcttcccc gcgcagggtc gcgatgctgc ccggtttggc actgctcctg 960
ctggccgcct ggacggctcg ggcgctggag gtacccactg atggtaatgc tggcctgctg 1020
gctgaacccc agattgccat gttctgtggc agactgaaca tgcacatgaa tgtccagaat 1080
gggaagtggg attcagatcc atcagggacc aaaacctgca ttgataccaa ggaaggcatc 1140
ctgcagtatt gccaagaagt ctaccctgaa ctgcagatca ccaatgtggt agaagccaac 1200
caaccagtga ccatccagaa ctggtgcaag cggggccgca agcagtgcaa gacccatccc 1260
cactttgtga ttccctaccg ctgcttagtt ggtgagttta taagtgatgc ccttctcgtt 1320
cctgacaagt gcaaattctt acaccaggag aggatggatg tttgcgaaac tcatcttcac 1380
tggcacaccg tcgccaaaga gacatgcagt gagaagagta ccaacttgca tgactacggc 1440
atgttgctgc cctgcggaat tgacaagttc cgaggggtag agtttgtgtg ttgcccactg 1500
gctgaagaaa gtgacaatgt ggattctgct gatgcggagg aggatgactc ggatgtctgg 1560
tggggcggag cagacacaga ctatgcagat gggagtgaag acaaagtagt agaagtagca 1620
gaggaggaag aagtggctga ggtggaagaa gaagaagccg atgatgacga ggacgatgag 1680
gatggtgatg aggtagagga agaggctgag gaaccctacg aagaagccac agagagaacc 1740
accagcattg ccaccaccac caccaccacc acagagtctg tggaagaggt ggttcgagtt 1800
cctacaacag cagccagtac ccctgatgcc gttgacaagt atctcgagac acctggggat 1860
gagaatgaac atgcccattt ccagaaagcc aaagagaggc ttgaggccaa gcaccgagag 1920
agaatgtccc aggtcatgag agaatgggaa gaggcagaac gtcaagcaaa gaacttgcct 1980
aaagctgata agaaggcagt tatccagcat ttccaggaga aagtggaatc tttggaacag 2040
gaagcagcca acgagagaca gcagctggtg gagacacaca tggccagagt ggaagccatg 2100
ctcaatgacc gccgccgcct ggccctggag aactacatca ccgctctgca ggctgttcct 2160
cctcggcctc gtcacgtgtt caatatgcta aagaagtatg tccgcgcaga acagaaggac 2220
agacagcaca ccctaaagca tttcgagcat gtgcgcatgg tggatcccaa gaaagccgct 2280
cagatccggt cccaggttat gacacacctc cgtgtgattt atgagcgcat gaatcagtct 2340
ctctccctgc tctacaacgt gcctgcagtg gccgaggaga ttcaggatga agttgatgag 2400
ctgcttcaga aagagcaaaa ctattcagat gacgtcttgg ccaacatgat tagtgaacca 2460
aggatcagtt acggaaacga tgctctcatg ccatctttga ccgaaacgaa aaccaccgtg 2520
gagctccttc ccgtgaatgg agagttcagc ctggacgatc tccagccgtg gcattctttt 2580
ggggctgact ctgtgccagc caacacagaa aacgaagttg agcctgttga tgcccgccct 2640
gctgccgacc gaggactgac cactcgacca ggttctgggt tgacaaatat caagacggag 2700
gagatctctg aagtgaatct agatgcagaa ttccgacatg actcaggata tgaagttcat 2760
catcaaaaat tggtgttctt tgcagaagat gtgggttcaa acaaaggtgc aatcattgga 2820
ctcatggtgg gcggtgttgt catagcgaca gtgatcgtca tcaccttggt gatgctgaag 2880
aagaaagata tcatggagcc agtagatcct agactagagc cctggaagca tccaggaagt 2940
cagcctaaaa ctgcttgtac caattgctat tgtaaaaagt gttgctttca ttgccaagtt 3000
tgtttcatga caaaagcctt aggcatctcc tatggcagga agaagcggag acagcgacga 3060
agagctcatc agaacagtca gactcatcaa gcttctctat caaagcagta agtaggcggc 3120
cgctcgagtc tagagggccc gtttaaaccc gctgatcagc ctcgactgtg ccttctagtt 3180
gccagccatc tgttgtttgc ccctcccccg tgccttcctt gaccctggaa ggtgccactc 3240
ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt 3300
ctattctggg gggtggggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca 3360
ggcatgctgg ggatgcggtg ggctctatgg cttctgaggc ggaaagaacc agctggggct 3420
ctaggggtta tccccacgcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta 3480
cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctcctttc gctttcttcc 3540
cttcctttct cgccacgttc gccggctttc cccgtcaagc tctaaatcgg ggcatcccct 3600

tagggttccg atttagtgct ttacggcacc tcgaccccaa aaaacttgat tagggtgatg 3660
gttcacgtag tgggccatcg ccctgataga cggtttttcg ccctttgacg ttggagtcca 3720
cgttctttaa tagtggactc ttgttccaaa ctggaacaac actcaaccct atctcggtct 3780
attcttttga tttataaggg attttgggga tttcggccta ttggttaaaa aatgagctga 3840
tttaacaaaa atttaacgcg aattaattct gtggaatgtg tgtcagttag ggtgtggaaa 3900
gtccccaggc tccccaggca ggcagaagta tgcaaagcat gcatctcaat tagtcagcaa 3960
ccaggtgtgg aaagtcccca ggctccccag caggcagaag tatgcaaagc atgcatctca 4020
attagtcagc aaccatagtc ccgcccctaa ctccgcccat cccgccccta actccgccca 4080
gttccgccca ttctccgccc catggctgac taatttttt tatttatgca gaggccgagg 4140
ccgcctctgc ctctgagcta ttccagaagt agtgaggagg ctttttgga ggcctaggct 4200
tttgcaaaaa gctcccggga gcttgtatat ccattttcgg atctgatcag cacgtgttga 4260
caattaatca tcggcatagt atatcggcat agtataatac gacaaggtga ggaactaaac 4320
catggccaag ttgaccagtg ccgttccggt gctcaccgcg cgcgacgtcg ccggagcggt 4380
cgagttctgg accgaccggc tcgggttctc ccgggacttc gtggaggacg acttcgccgg 4440
tgtggtccgg gacgacgtga ccctgttcat cagcgcggtc caggaccagg tggtgccgga 4500
caacaccctg gcctgggtgt gggtgcgcgg cctggacgag ctgtacgccg agtggtcgga 4560
ggtcgtgtcc acgaacttcc gggacgcctc cgggccggcc atgaccgaga tcggcgagca 4620
gccgtggggg cgggagttcg ccctgcgcga cccggccggc aactgcgtgc acttcgtggc 4680
cgaggagcag gactgacacg tgctacgaga tttcgattcc accgccgcct tctatgaaag 4740
gttgggcttc ggaatcgttt tccgggacgc cggctggatg atcctccagc gcggggatct 4800
catgctggag ttcttcgccc accccaactt gtttattgca gcttataatg gttacaaata 4860
aagcaatagc atcacaaatt tcacaaataa agcatttttt tcactgcatt ctagttgtgg 4920
tttgtccaaa ctcatcaatg tatcttatca tgtctgtata ccgtcgacct ctagctagag 4980
cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc 5040
acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta 5100
actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca 5160
gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc 5220
cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc 5280
tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat 5340
gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 5400
ccataggctc cgcccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg 5460
aaacccgaca ggactataaa gataccaggc gtttcccccc ggaagctccc tcgtgcgctc 5520
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 5580
ggcgctttct caatgctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa 5640
gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta 5700
tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa 5760
caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa 5820
ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt 5880
cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt 5940
ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat 6000
cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat 6060
gagattatca aaaaggatct tcacctagat cctttttaaat taaaaatgaa gttttaaatc 6120
aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc 6180
acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta 6240
gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga 6300
cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg 6360
cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc 6420
tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat 6480
cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag 6540
gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat 6600
cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa 6660
ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa 6720
gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga 6780
taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg 6840
gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc 6900
acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg 6960
aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact 7020
cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat 7080
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 7140
gccacctgac gtc 7153

<210> 22
<211> 7153
<212> DNA
<213> human

<400> 22
ctgcctagcc ctctagaggg ctaggggata ccagctgaga gtcatgttag acgagactac 60
ggcgtatcaa ttcggtcata gacgagggac gaacacacaa cctccagcga ctcatcacgc 120
gctcgtttta aattcgatgt tgttccgttc cgaactggct gttaacgtac ttcttagacg 180
aatcccaatc cgcaaaacgc gacgaagcgc tacatgcccg gtctatatgc gcaactgtaa 240
ctaataactg atcaataatt atcattagtt aatgccccag taatcaagta tcgggtatat 300
acctcaaggc gcaatgtatt gaatgccatt taccgggcgg accgactggc gggttgctgg 360
gggcgggtaa ctgcagttat tactgcatac aagggtatca ttgcggttat ccctgaaagg 420
taactgcagt tacccacctg ataaatgcca tttgacgggt gaaccgtcat gtagttcaca 480
tagtatacgg ttcatgcggg ggataactgc agttactgcc atttaccggg cggaccgtaa 540
tacgggtcat gtactggaat accctgaaag gatgaaccgt catgtagatg cataatcagt 600
agcgataatg gtaccactac gccaaaaccg tcatgtagtt acccgcacct atcgccaaac 660
tgagtgcccc taaaggttca gaggtggggt aactgcagtt accctcaaac aaaaccgtgg 720
ttttagttgc cctgaaaggt tttacagcat tgttgaggcg gggtaactgc gtttacccgc 780
catccgcaca tgccaccctc cagatatatt cgtctcgaga gaccgattga tctcttgggt 840
gacgaatgac cgaatagctt taattatgct gagtgatatc cctctgggtt cgaccgatcg 900
caaatttgaa ttcgaagggg cgcgtcccag cgctacgacg ggccaaaccg tgacgaggac 960
gaccggcgga cctgccgagc ccgcgacctc catgggtgac taccattacg accggacgac 1020
cgacttgggg tctaacggta caagacaccg tctgacttgt acgtgtactt acaggtctta 1080
cccttcaccc taagtctagg tagtccctgg ttttggacgt aactatggtt ccttccgtag 1140
gacgtcataa cggttcttca gatgggactt gacgtctagt ggttacacca tcttcggttg 1200
gttggtcact ggtaggtctt gaccacgttc gccccggcgt tcgtcacgtt ctgggtaggg 1260
gtgaaacact aagggatggc gacgaatcaa ccactcaaat attcactacg ggaagagcaa 1320
ggactgttca cgtttaagaa tgtggtcctc tcctacctac aaacgctttg agtagaagtg 1380
accgtgtggc agcggtttct ctgtacgtca ctcttctcat ggttgaacgt actgatgccg 1440
tacaacgacg ggacgcctta actgttcaag gctccccatc tcaaacacac aacgggtgac 1500
cgacttcttt cactgttaca cctaagacga ctacgcctcc tcctactgag cctacagacc 1560
accccgcctc gtctgtgtct gatacgtcta ccctcacttc tgtttcatca tcttcatcgt 1620
ctcctccttc ttcaccgact ccaccttctt cttcttcggc tactactgct cctgctactc 1680
ctaccactac tccatctcct tctccgactc cttgggatgc ttcttcggtg tctctcttgg 1740
tggtcgtaac ggtggtggtg gtggtggtgg tgtctcagac accttctcca ccaagctcaa 1800
ggatgttgtc gtcggtcatg gggactacgg caactgttca tagagctctg tggaccccta 1860
ctcttacttg tacgggtaaa ggtctttcgg tttctctccg aactccggtt cgtggctctc 1920
tcttacaggg tccagtactc tcttaccctt ctccgtcttg cagttcgttt cttgaacgga 1980
tttcgactat tcttccgtca ataggtcgta aaggtcctct ttcaccttag aaaccttgtc 2040
cttcgtcggt tgctctctgt cgtcgaccac ctctgtgtgt accggtctca ccttcggtac 2100
gagttactgg cggcggcgga ccgggacctc ttgatgtagt ggcgagacgt ccgacaagga 2160
ggagccggag cagtgcacaa gttatacgat ttcttcatac aggcgcgtct tgtcttcctg 2220
tctgtcgtgt gggatttcgt aaagctcgta cacgcgtacc acctagggtt ctttcggcga 2280
gtctaggcca gggtccaata ctgtgtggag gcacactaaa tactcgcgta cttagtcaga 2340
gagagggacg agatgttgca cggacgtcac cggctcctct aagtcctact tcaactactc 2400
gacgaagtct ttctcgtttt gataagtcta ctgcagaacc ggttgtacta atcacttggt 2460
tcctagtcaa tgcctttgct acgagagtac ggtagaaact ggctttgctt ttggtggcac 2520
ctcgaggaag ggcacttacc tctcaagtcg gacctgctag aggtcggcac cgtaagaaaa 2580
ccccgactga gacacggtcg gttgtgtctt ttgcttcaac tcggacaact acgggcggga 2640
cgacggctgg ctcctgactg gtgagctggt ccaagaccca actgtttata gttctgcctc 2700
ctctagagac ttcacttaga tctacgtctt aaggctgtac tgagtcctat acttcaagta 2760
gtagttttta accacaagaa acgtcttcta cacccaagtt tgtttccacg ttagtaacct 2820
gagtaccacc cgccacaaca gtatcgctgt cactagcagt agtggaacca ctacgacttc 2880
ttctttctat agtacctcgg tcatctagga tctgatctcg ggaccttcgt aggtccttca 2940
gtcggatttt gacgaacatg gttaacgata acatttttca caacgaaagt aacggttcaa 3000
acaaagtact gttttcggaa tccgtagagg ataccgtcct tcttcgcctc tgtcgctgct 3060
tctcgagtag tcttgtcagt ctgagtagtt cgaagagata gtttcgtcat tcatccgccg 3120
gcgagctcag atctcccggg caaatttggg cgactagtcg gagctgacac ggaagatcaa 3180
cggtcggtag acaacaaacg gggagggggc acggaaggaa ctgggacctt ccacggtgag 3240
ggtgacagga aaggattatt ttactccttt aacgtagcgt aacagactca tccacagtaa 3300

gataagaccc cccaccccac cccgtcctgt cgttcccccc cctaaccctt ctgttatcgt 3360
ccgtacgacc cctacgccac ccgagatacc gaagactccg cctttcttgg tcgaccccga 3420
gatcccccat aggggtgcgc gggacatcgc cgcgtaattc gcgccgccca caccaccaat 3480
gcgcgtcgca ctggcgatgt gaacggtcgc gggatcgcgg gcgaggaaag cgaaagaagg 3540
gaaggaaaga gcggtgcaag cggccgaaag gggcagttcg agatttagcc ccgtagggaa 3600
atcccaaggc taaatcacga aatgccgtgg agctggggtt ttttgaacta atcccactac 3660
caagtgcatc acccggtagc gggactatct gccaaaaagc gggaaactgc aacctcaggt 3720
gcaagaaatt atcacctgag aacaaggttt gaccttgttg tgagttggga tagagccaga 3780
taagaaaact aaatattccc taaaacccct aaagccggat aaccaatttt ttactcgact 3840
aaattgtttt taaattgcgc ttaattaaga caccttacac acagtcaatc ccacaccttt 3900
caggggtccg aggggtccgt ccgtcttcat acgtttcgta cgtagagtta atcagtcgtt 3960
ggtccacacc tttcaggggt ccgaggggtc gtccgtcttc atacgtttcg tacgtagagt 4020
taatcagtcg ttggtatcag ggcggggatt gaggcgggta gggcggggat tgaggcgggt 4080
caaggcgggt aagaggcggg gtaccgactg attaaaaaaa ataaatacgt ctccggctcc 4140
ggcggagacg gagactcgat aaggtcttca tcactcctcc gaaaaaacct ccggatccga 4200
aaacgttttt cgagggccct cgaacatata ggtaaaagcc tagactagtc gtgcacaact 4260
gttaattagt agccgtatca tatagccgta tcatattatg ctgttccact ccttgatttg 4320
gtaccggttc aactggtcac ggcaaggcca cgagtggcgc gcgctgcagc ggcctcgcca 4380
gctcaagacc tggctggccg agcccaagag ggccctgaag cacctcctgc tgaagcggcc 4440
acaccaggcc ctgctgcact gggacaagta gtcgcgccag gtcctggtcc accacggcct 4500
gttgtgggac cggacccaca cccacgcgcc ggacctgctc gacatgcggc tcaccagcct 4560
ccagcacagg tgcttgaagg ccctgcggag gcccggccgg tactggctct agccgctcgt 4620
cggcaccccc gccctcaagc gggacgcgct gggccggccg ttgacgcacg tgaagcaccg 4680
gctcctcgtc ctgactgtgc acgatgctct aaagctaagg tggcggcgga agatactttc 4740
caacccgaag ccttagcaaa aggccctgcg gccgacctac taggaggtcg cgcccctaga 4800
gtacgacctc aagaagcggg tggggttgaa caaataacgt cgaatattac caatgtttat 4860
ttcgttatcg tagtgtttaa agtgtttatt tcgtaaaaaa agtgacgtaa gatcaacacc 4920
aaacaggttt gagtagttac atagaatagt acagacatat ggcagctgga gatcgatctc 4980
gaaccgcatt agtaccagta tcgacaaagg acacacttta acaataggcg agtgttaagg 5040
tgtgttgtat gctcggcctt cgtatttcac atttcggacc ccacggatta ctcactcgat 5100
tgagtgtaat taacgcaacg cgagtgacgg gcgaaaggtc agcccttggg acagcacggt 5160
cgacgtaatt acttagccgg ttgcgcgccc ctctccgcca aacgcataac ccgcgagaag 5220
gcgaaggagc gagtgactga gcgacgcgag ccagcaagcc gacgccgctc gccatagtcg 5280
agtgagtttc cgccattatg ccaataggtg tcttagtccc ctattgcgtc ctttcttgta 5340
cactcgtttt ccggtcgttt tccggtcctt ggcatttttc cggcgcaacg accgcaaaaa 5400
ggtatccgag gcgggggggac tgctcgtagt gtttttagct gcgagttcag tctccaccgc 5460
tttgggctgt cctgatattt ctatggtccg caaaggggga ccttcgaggg agcacgcgag 5520
aggacaaggc tgggacggcg aatggcctat ggacaggcgg aaagagggaa gcccttcgca 5580
ccgcgaaaga gttacgagtg cgacatccat agagtcaagc cacatccagc aagcgaggtt 5640
cgacccgaca cacgtgcttg gggggcaagt cgggctggcg acgcggaata ggccattgat 5700
agcagaactc aggttgggcc attctgtgct gaatagcggt gaccgtcgtc ggtgaccatt 5760
gtcctaatcg tctcgctcca tacatccgcc acgatgtctc aagaacttca ccaccggatt 5820
gatgccgatg tgatcttcct gtcataaacc atagacgcga gacgacttcg gtcaatggaa 5880
gccttttttct caaccatcga gaactaggcc gtttgtttgg tggcgaccat cgccaccaaa 5940
aaaacaaacg ttcgtcgtct aatgcgcgtc ttttttttcct agagttcttc taggaaacta 6000
gaaaagatgc cccagactgc gagtcacctt gcttttgagt gcaattccct aaaaccagta 6060
ctctaatagt ttttcctaga agtggatcta ggaaaattta atttttactt caaaatttag 6120
ttagatttca tatatactca tttgaaccag actgtcaatg gttacgaatt agtcactccg 6180
tggatagagt cgctagacag ataaagcaag taggtatcaa cggactgagg ggcagcacat 6240
ctattgatgc tatgccctcc cgaatggtag accggggtca cgacgttact atggcgctct 6300
gggtgcgagt ggccgaggtc taaatagtcg ttatttggtc ggtcggcctt cccggctcgc 6360
gtcttcacca ggacgttgaa ataggcggag gtaggtcaga taattaacaa cggcccttcg 6420
atctcattca tcaagcggtc aattatcaaa cgcgttgcaa caacggtaac gatgtccgta 6480
gcaccacagt gcgagcagca aaccataccg aagtaagtcg aggccaaggg ttgctagttc 6540
cgctcaatgt actagggggt acaacacgtt ttttcgccaa tcgaggaagc caggaggcta 6600
gcaacagtct tcattcaacc ggcgtcacaa tagtgagtac caataccgtc gtgacgtatt 6660
aagagaatga cagtacggta ggcattctac gaaaagacac tgaccactca tgagttggtt 6720
cagtaagact cttatcacat acgccgctgg ctcaacgaga acgggccgca gttatgccct 6780
attatggcgc ggtgtatcgt cttgaaattt tcacgagtag taacctttg caagaagccc 6840
cgcttttgag agttcctaga atggcgacaa ctctaggtca agctacattg ggtgagcacg 6900
tgggttgact agaagtcgta gaaaatgaaa gtggtcgcaa agacccactc gtttttgtcc 6960
ttccgtttta cggcgttttt tcccttattc ccgctgtgcc tttacaactt atgagtatga 7020

gaaggaaaaa gttataataa cttcgtaaat agtcccaata acagagtact cgcctatgta 7080
taaacttaca taaatctttt tatttgttta tccccaaggc gcgtgtaaag gggcttttca 7140
cggtggactg cag 7153

<210> 23
<211> 2352
<212> DNA
<213> fusion protein - human

<400> 23
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaactttga agtggaattc 1800
cgacatgact caggatatga agttcatcat caagtattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc ctggagccag tagatcctag actagagccc 2040
tggaagcatc caggaagtca gcctaaaact gcttgtacca attgctattg taaaaagtgt 2100
tgctttcatt gccaagtttg tttcatgaca aaagccttag gcatctccta tggcaggaag 2160
aagcggagac agcgacgaag agctcatcag aacagtcaga ctcatcaagc ttctctatca 2220
aagcagagga tatccagcac agtggcggcc gcagacgccg ctgtcacccc agaggagcgc 2280
cacctgtcca agatgcagca gaacggctac gaaaatccaa cctacaagtt ctttgagcag 2340
atgcagaact ag 2352

<210> 24
<211> 783
<212> PRT
<213> fusion protein - human

<400> 24
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro

```
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
```

```
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
            660                 665                 670
Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
        675                 680                 685
Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
    690                 695                 700
Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
705                 710                 715                 720
Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr His Gln
                725                 730                 735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Ala Asp
            740                 745                 750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
        755                 760                 765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    770                 775                 780
```

<210> 25
<211> 2823
<212> DNA
<213> fusion protein - human

<400> 25

```
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
```

```
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaactttga agtggaattc 1800
cgacatgact caggatatga agttcatcat caagtattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc aagctactgt cttctatcga acaagcatgc 2040
gatatttgcc gacttaaaaa gctcaagtgc tccaaagaaa aaccgaagtg cgccaagtgt 2100
ctgaagaaca actgggagtg tcgctactct cccaaaacca aaaggtctcc gctgactagg 2160
gcacatctga cagaagtgga atcaaggcta gaaagactgg aacagctatt tctactgatt 2220
tttcctcgag aagaccttga catgattttg aaaatggatt ctttacagga tataaaagca 2280
ttgttaacag gattatttgt acaagataat gtgaataaag atgccgtcac agatagattg 2340
gcttcagtgg agactgatat gcctctaaca ttgagacagc atagaataag tgcgacatca 2400
tcatcggaag agagtagtaa caaaggtcaa agacagttga ctgtatcggg aattcccggg 2460
gatctggccc ccccgaccga tgtcagcctg ggggacgagc tccacttaga cggcgaggac 2520
gtggcgatgg cgcatgccga cgcgctagac gatttcgatc tggacatgtt gggggacggg 2580
gattccccgg gtccgggatt tacccccccac gactccgccc cctacggcgc tctggatatg 2640
gccgacttcg agtttgagca gatgtttacc gatgcccttg gaattgacga gtacggtggg 2700
gatatccagc acagtggcgg ccgcgacgcc gctgtcaccc cagaggagcg ccacctgtcc 2760
aagatgcagc agaacggcta cgaaaatcca acctacaagt tctttgagca gatgcagaac 2820
tag                                                               2823
```

<210> 26
<211> 941
<212> PRT
<213> fusion protein - human

<400> 26

```
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1               5                  10                  15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20                  25                  30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
```

```
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
            660                 665                 670
Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
        675                 680                 685
Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
    690                 695                 700
Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
705                 710                 715                 720
```

```
Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
                725                 730                 735
Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
            740                 745                 750
Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
        755                 760                 765
Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
    770                 775                 780
Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
785                 790                 795                 800
Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
                805                 810                 815
Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
            820                 825                 830
Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
        835                 840                 845
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
    850                 855                 860
Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
865                 870                 875                 880
Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
                885                 890                 895
Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala
            900                 905                 910
Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
        915                 920                 925
Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    930                 935                 940
```

<210> 27
<211> 5015
<212> DNA
<213> human

<400> 27

```
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc 900
gtttaaactt aagcttggta ccgagctcgg atccactagt ccagtgtggt ggaattctgc 960
agatatccag cacagtggcg gccgctcgag tctagagggc ccgtttaaac ccgctgatca 1020
gcctcgactg tgccttctag ttgccagcca tctgttgttt gcccctcccc cgtgccttcc 1080
ttgaccctgg aaggtgccac tcccactgtc ctttcctaat aaaatgagga aattgcatcg 1140
cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg 1200
gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat ggcttctgag 1260
gcggaaagaa ccagctgggg ctctaggggg tatccccacg cgccctgtag cggcgcatta 1320
agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg 1380
cccgctcctt tcgctttctt cccttcctttctcgccacgt tcgccggctt tccccgtcaa 1440
gctctaaatc ggggcatccc tttagggttc cgatttagtg ctttacggca cctcgacccc 1500
aaaaaacttg attagggtga tggttcacgt agtgggccat cgccctgata gacggttttt 1560
```

cgccctttga cgttggagtc cacgttcttt aatagtggac tcttgttcca aactggaaca 1620
acactcaacc ctatctcggt ctattctttt gatttataag ggattttggg gatttcggcc 1680
tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattaatt ctgtggaatg 1740
tgtgtcagtt agggtgtgga aagtccccag gctccccagg caggcagaag tatgcaaagc 1800
atgcatctca attagtcagc aaccaggtgt ggaaagtccc caggctcccc agcaggcaga 1860
agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgccccc aactccgccc 1920
atcccgcccc taactccgcc cagttccgcc cattctccgc cccatggctg actaatttt 1980
tttatttatg cagaggccga ggccgcctct gcctctgagc tattccagaa gtagtgagga 2040
ggctttttg gaggcctagg cttttgcaaa aagctcccgg gagcttgtat atccattttc 2100
ggatctgatc agcacgtgtt gacaattaat catcggcata gtatatcggc atagtataat 2160
acgacaaggt gaggaactaa accatggcca agttgaccag tgccgttccg gtgctcaccg 2220
cgcgcgacgt cgccggagcg gtcgagttct ggaccgaccg gctcgggttc tcccgggact 2280
tcgtggagga cgacttcgcc ggtgtggtcc gggacgacgt gaccctgttc atcagcgcgg 2340
tccaggacca ggtggtgccg gacaacaccc tggcctgggt gtgggtgcgc ggcctggacg 2400
agctgtacgc cgagtggtcg gaggtcgtgt ccacgaactt ccgggacgcc tccgggccgg 2460
ccatgaccga gatcggcgag cagccgtggg ggcgggagtt cgccctgcgc gacccggccg 2520
gcaactgcgt gcacttcgtg gccgaggagc aggactgaca cgtgctacga gatttcgatt 2580
ccaccgccgc cttctatgaa aggttgggct tcggaatcgt tttccgggac gccggctgga 2640
tgatcctcca gcgcggggat ctcatgctgg agttcttcgc ccaccccaac ttgtttattg 2700
cagcttataa tggttacaaa taaagcaata gcatcacaaa tttcacaaat aaagcatttt 2760
tttcactgca ttctagttgt ggtttgtcca aactcatcaa tgtatcttat catgtctgta 2820
taccgtcgac ctctagctag agcttggcgt aatcatggtc atagctgttt cctgtgtgaa 2880
attgttatcc gctcacaatt ccacacaaca tacgagccgg aagcataaag tgtaaagcct 2940
ggggtgccta atgagtgagc taactcacat taattgcgtt gcgctcactg cccgctttcc 3000
agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg 3060
gtttgcgtat tgggcgctct tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc 3120
ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag 3180
gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa 3240
aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc 3300
gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc 3360
ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg 3420
cctttctccc ttcgggaagc gtggcgcttt ctcaatgctc acgctgtagg tatctcagtt 3480
cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccgtt cagcccgacc 3540
gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc 3600
cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 3660
agttcttgaa gtggtggcct aactacggct acactagaag gacagtattt ggtatctgcg 3720
ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa 3780
ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag 3840
gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact 3900
cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa 3960
attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt 4020
accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag 4080
ttgcctgact ccccgtcgtg tagataacta cgatacggga gggcttacca tctggcccca 4140
gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca gcaataaacc 4200
agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt 4260
ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg 4320
ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg gcttcattca 4380
gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg 4440
ttagctcctt cggtcctccg atcgttgtca gaagtaagtt ggccgcagtg ttatcactca 4500
tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg 4560
tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct 4620
cttgcccggc gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgctca 4680
tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca 4740
gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg 4800
tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac 4860
ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt 4920
attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa atagggggttc 4980
cgcgcacatt tccccgaaaa gtgccacctg acgtc 5015

<210> 28

<211> 5015
<212> DNA
<213> human

<400> 28
ctgcctagcc ctctagaggg ctaggggata ccagctgaga gtcatgttag acgagactac 60
ggcgtatcaa ttcggtcata gacgagggac gaacacacaa cctccagcga ctcatcacgc 120
gctcgtttta aattcgatgt tgttccgttc cgaactggct gttaacgtac ttcttagacg 180
aatcccaatc cgcaaaacgc gacgaagcgc tacatgcccg gtctatatgc gcaactgtaa 240
ctaataactg atcaataatt atcattagtt aatgccccag taatcaagta tcgggtatat 300
acctcaaggc gcaatgtatt gaatgccatt taccgggcgg accgactggc gggttgctgg 360
gggcgggtaa ctgcagttat tactgcatac aagggtatca ttgcggttat ccctgaaagg 420
taactgcagt tacccacctg ataaatgcca tttgacgggt gaaccgtcat gtagttcaca 480
tagtatacgg ttcatgcggg ggataactgc agttactgcc atttaccggg cggaccgtaa 540
tacgggtcat gtactggaat accctgaaag gatgaaccgt catgtagatg cataatcagt 600
agcgataatg gtaccactac gccaaaaccg tcatgtagtt acccgcacct atcgccaaac 660
tgagtgcccc taaaggttca gaggtggggt aactgcagtt accctcaaac aaaaccgtgg 720
ttttagttgc cctgaaaggt tttacagcat tgttgaggcg gggtaactgc gtttacccgc 780
catccgcaca tgccaccctc cagatatatt cgtctcgaga gaccgattga tctcttgggt 840
gacgaatgac cgaatagctt taattatgct gagtgatatc cctctgggtt cgaccgatcg 900
caaatttgaa ttcgaaccat ggctcgagcc taggtgatca ggtcacacca ccttaagacg 960
tctataggtc gtgtcaccgc cggcgagctc agatctcccg ggcaaatttg ggcgactagt 1020
cggagctgac acggaagatc aacggtcggt agacaacaaa cggggagggg gcacggaagg 1080
aactgggacc ttccacggtg agggtgacag gaaaggatta ttttactcct ttaacgtagc 1140
gtaacagact catccacagt aagataagac cccccacccc accccgtcct gtcgttcccc 1200
ctcctaaccc ttctgttatc gtccgtacga cccctacgcc acccgagata ccgaagactc 1260
cgcctttctt ggtcgacccc gagatccccc ataggggtgc gcgggacatc gccgcgtaat 1320
tcgcgccgcc cacaccacca atgcgcgtcg cactggcgat gtgaacggtc gcgggatcgc 1380
gggcgaggaa agcgaaagaa gggaaggaaa gagcggtgca agcggccgaa aggggcagtt 1440
cgagatttag ccccgtaggg aaatcccaag gctaaatcac gaaatgccgt ggagctgggg 1500
ttttttgaac taatcccact accaagtgca tcacccggta gcgggactat ctgccaaaaa 1560
gcgggaaact gcaacctcag gtgcaagaaa ttatcacctg agaacaaggt ttgaccttgt 1620
tgtgagttgg gatagagcca gataagaaaa ctaaatattc cctaaaaccc ctaaagccgg 1680
ataaccaatt ttttactcga ctaaattgtt tttaaattgc gcttaattaa gacaccttac 1740
acacagtcaa tcccacacct ttcaggggtc cgaggggtcc gtccgtcttc atacgtttcg 1800
tacgtagagt taatcagtcg ttggtccaca cctttcaggg gtccgagggg tcgtccgtct 1860
tcatacgttt cgtacgtaga gttaatcagt cgttggtatc agggcgggga ttgaggcggg 1920
tagggcgggg attgaggcgg gtcaaggcgg gtaagaggcg gggtaccgac tgattaaaaa 1980
aaataaatac gtctccggct ccggcggaga cggagactcg ataaggtctt catcactcct 2040
ccgaaaaaac ctccggatcc gaaaacgttt ttcgagggcc ctcgaacata taggtaaaag 2100
cctagactag tcgtgcacaa ctgttaatta gtagccgtat catatagccg tatcatatta 2160
tgctgttcca ctccttgatt tggtaccggt tcaactggtc acggcaaggc cacgagtggc 2220
gcgcgctgca gcggcctcgc cagctcaaga cctggctggc cgagcccaag agggccctga 2280
agcacctcct gctgaagcgg ccacaccagg ccctgctgca ctgggacaag tagtcgcgcc 2340
aggtcctggt ccaccacggc ctgttgtggg accggaccca cacccacgcg ccggacctgc 2400
tcgacatgcg gctcaccagc ctccagcaca ggtgcttgaa ggccctgcgg aggcccggcc 2460
ggtactggct ctagccgctc gtcggcaccc ccgccctcaa gcgggacgcg ctgggccggc 2520
cgttgacgca cgtgaagcac cggctcctcg tcctgactgt gcacgatgct ctaaagctaa 2580
ggtggcggcg gaagatactt tccaacccga agccttagca aaaggccctg cggccgacct 2640
actaggaggt cgcgccccta gagtacgacc tcaagaagcg ggtggggttg aacaaataac 2700
gtcgaatatt accaatgttt atttcgttat cgtagtgttt aaagtgttta tttcgtaaaa 2760
aaagtgacgt aagatcaaca ccaaacaggt ttgagtagtt acatagaata gtacagacat 2820
atggcagctg gagatcgatc tcgaaccgca ttagtaccag tatcgacaaa ggacacactt 2880
taacaatagg cgagtgttaa ggtgtgttgt atgctcggcc ttcgtatttc acatttcgga 2940
ccccacggat tactcactcg attgagtgta attaacgcaa cgcgagtgac gggcgaaagg 3000
tcagcccttt ggacagcacg gtcgacgtaa ttacttagcc ggttgcgcgc ccctctccgc 3060
caaacgcata acccgcgaga aggcgaagga gcgagtgact gagcgacgcg agccagcaag 3120
ccgacgccgc tcgccatagt cgagtgagtt tccgccatta tgccaatagg tgtcttagtc 3180
ccctattgcg tcctttcttg tacactcgtt ttccggtcgt tttccggtcc ttggcatttt 3240
tccggcgcaa cgaccgcaaa aaggtatccg aggcgggggg actgctcgta gtgtttttag 3300
ctgcgagttc agtctccacc gctttgggct gtcctgatat ttctatggtc cgcaaagggg 3360
gaccttcgag ggagcacgcg agaggacaag gctgggacgg cgaatggcct atggacaggc 3420

ggaaagaggg aagcccttcg caccgcgaaa gagttacgag tgcgacatcc atagagtcaa 3480
gccacatcca gcaagcgagg ttcgacccga cacacgtgct tgggggcaa gtcgggctgg 3540
cgacgcggaa taggccattg atagcagaac tcaggttggg ccattctgtg ctgaatagcg 3600
gtgaccgtcg tcggtgacca ttgtcctaat cgtctcgctc catacatccg ccacgatgtc 3660
tcaagaactt caccaccgga ttgatgccga tgtgatcttc ctgtcataaa ccatagacgc 3720
gagacgactt cggtcaatgg aagccttttt ctcaaccatc gagaactagg ccgtttgttt 3780
ggtggcgacc atcgccacca aaaaaacaaa cgttcgtcgt ctaatgcgcg tctttttttc 3840
ctagagttct tctaggaaac tagaaaagat gccccagact gcgagtcacc ttgcttttga 3900
gtgcaattcc ctaaaaccag tactctaata gtttttccta gaagtggatc taggaaaatt 3960
taatttttac ttcaaaattt agttagattt catatatact catttgaacc agactgtcaa 4020
tggttacgaa ttagtcactc cgtggataga gtcgctagac agataaagca agtaggtatc 4080
aacggactga ggggcagcac atctattgat gctatgccct cccgaatggt agaccggggt 4140
cacgacgtta ctatggcgct ctgggtgcga gtggccgagg tctaaatagt cgttatttgg 4200
tcggtcggcc ttcccggctc gcgtcttcac caggacgttg aaataggcgg aggtaggtca 4260
gataattaac aacggccctt cgatctcatt catcaagcgg tcaattatca aacgcgttgc 4320
aacaacggta acgatgtccg tagcaccaca gtgcgagcag caaaccatac cgaagtaagt 4380
cgaggccaag ggttgctagt tccgctcaat gtactagggg gtacaacacg ttttttcgcc 4440
aatcgaggaa gccaggaggc tagcaacagt cttcattcaa ccggcgtcac aatagtgagt 4500
accaataccg tcgtgacgta ttaagagaat gacagtacgg taggcattct acgaaaagac 4560
actgaccact catgagttgg ttcagtaaga ctcttatcac atacgccgct ggctcaacga 4620
gaacgggccg cagttatgcc ctattatggc gcggtgtatc gtcttgaaat tttcacgagt 4680
agtaaccttt tgcaagaagc cccgcttttg agagttccta gaatggcgac aactctaggt 4740
caagctacat tgggtgagca cgtgggttga ctagaagtcg tagaaaatga aagtggtcgc 4800
aaagacccac tcgtttttgt ccttccgttt tacggcgttt tttcccttat tcccgctgtg 4860
cctttacaac ttatgagtat gagaaggaaa aagttataat aacttcgtaa atagtcccaa 4920
taacagagta ctcgcctatg tataaactta cataaatctt tttatttgtt tatccccaag 4980
gcgcgtgtaa aggggctttt cacggtggac tgcag 5015

<210> 29
<211> 2352
<212> DNA
<213> fusion protein - human

<400> 29
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560

tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaactttga agtggaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc ctggagccag tagatcctag actagagccc 2040
tggaagcatc caggaagtca gcctaaaact gcttgtacca attgctattg taaaaagtgt 2100
tgctttcatt gccaagtttg tttcatgaca aaagccttag gcatctccta tggcaggaag 2160
aagcggagac agcgacgaag agctcatcag aacagtcaga ctcatcaagc ttctctatca 2220
aagcagagga tatccagcac agtggcggcc gcagacgccg ctgtcacccc agaggagcgc 2280
cacctgtcca agatgcagca gaacggctac gaaaatccaa cctacaagtt ctttgagcag 2340
atgcagaact ag 2352

<210> 30
<211> 783
<212> PRT
<213> fusion protein - human

<400> 30
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
115 120 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145 150 155 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
165 170 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
180 185 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
195 200 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
210 215 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225 230 235 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
245 250 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
260 265 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
275 280 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
290 295 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305 310 315 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg

325 330 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
340 345 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
355 360 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
370 375 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385 390 395 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
405 410 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
420 425 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
435 440 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
450 455 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465 470 475 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
485 490 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
500 505 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
515 520 525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
530 535 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545 550 555 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
565 570 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
580 585 590
Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser Gly Tyr Glu Val
595 600 605
His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
610 615 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625 630 635 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
645 650 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
660 665 670
Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
675 680 685
Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
690 695 700
Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
705 710 715 720
Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr His Gln
725 730 735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Ala Asp
740 745 750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
755 760 765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
770 775 780

<210> 31
<211> 2823
<212> DNA

<213> fusion protein - human

<400> 31
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttataa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaactttga agtggaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaaaagcttg gtaccgagct cggatccact 1980
agtccagtgt ggtggaattc tgcagatatc aagctactgt cttctatcga acaagcatgc 2040
gatatttgcc gacttaaaaa gctcaagtgc tccaaagaaa aaccgaagtg cgccaagtgt 2100
ctgaagaaca actgggagtg tcgctactct cccaaaacca aaaggtctcc gctgactagg 2160
gcacatctga cagaagtgga atcaaggcta gaaagactgg aacagctatt tctactgatt 2220
tttcctcgag aagaccttga catgattttg aaaatggatt ctttacagga tataaaagca 2280
ttgttaacag gattatttgt acaagataat gtgaataaag atgccgtcac agatagattg 2340
gcttcagtgg agactgatat gcctctaaca ttgagacagc atagaataag tgcgacatca 2400
tcatcggaag agagtagtaa caaaggtcaa agacagttga ctgtatcggg aattcccggg 2460
gatctggccc ccccgaccga tgtcagcctg ggggacgagc tccacttaga cggcgaggac 2520
gtggcgatgg cgcatgccga cgcgctagac gatttcgatc tggacatgtt gggggacggg 2580
gattccccgg gtccgggatt tacccccac gactccgccc cctacggcgc tctggatatg 2640
gccgacttcg agtttgagca gatgtttacc gatgcccttg gaattgacga gtacggtggg 2700
gatatccagc acagtggcgg ccgcgacgcc gctgtcaccc cagaggagcg ccacctgtcc 2760
aagatgcagc agaacggcta cgaaaatcca acctacaagt tctttgagca gatgcagaac 2820
tag 2823

<210> 32
<211> 941
<212> PRT
<213> fusion protein - human

<400> 32
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1   5   10   15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20   25   30

```
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35                  40                  45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50                  55                  60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65                  70                  75                  80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85                  90                  95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
        115                 120                 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
        355                 360                 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525
```

```
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605
His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
                645                 650                 655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
            660                 665                 670
Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
        675                 680                 685
Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
    690                 695                 700
Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
705                 710                 715                 720
Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
                725                 730                 735
Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
            740                 745                 750
Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
        755                 760                 765
Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
    770                 775                 780
Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
785                 790                 795                 800
Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
                805                 810                 815
Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
            820                 825                 830
Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
        835                 840                 845
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
    850                 855                 860
Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
865                 870                 875                 880
Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
                885                 890                 895
Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala
            900                 905                 910
Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
        915                 920                 925
Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
    930                 935                 940
```

<210> 33
<211> 63
<212> PRT
<213> human

<400> 33
Glu Glu Ile Ser Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser
 1   5   10   15

```
Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val
            20                  25                  30
Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val
        35                  40                  45
Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys
    50                  55                  60


<210> 34
<211> 4
<212> PRT
<213> human

<400> 34
Lys Met Asp Ala
 1


<210> 35
<211> 7
<212> PRT
<213> human

<400> 35
Gly Val Val Ile Ala Thr Val
 1               5


<210> 36
<211> 63
<212> PRT
<213> human

<400> 36
Glu Glu Ile Ser Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser
 1               5                  10                  15
Gly Tyr Glu Val His His Gln Val Leu Val Phe Phe Ala Glu Asp Val
            20                  25                  30
Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val
        35                  40                  45
Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys
    50                  55                  60


<210> 37
<211> 63
<212> PRT
<213> human

<400> 37
Glu Glu Ile Ser Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser
 1               5                  10                  15
Gly Tyr Glu Val His His Gln Val Leu Val Phe Phe Ala Glu Asp Val
            20                  25                  30
Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val
        35                  40                  45
Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys
    50                  55                  60


<210> 38
<211> 4
```

<212> PRT
<213> human

<400> 38
Asn Leu Asp Ala
1

<210> 39
<211> 63
<212> PRT
<213> human

<400> 39
Glu Glu Ile Ser Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser
1 5 10 15
Gly Tyr Glu Val His His Gln Val Leu Val Phe Phe Ala Glu Asp Val
20 25 30
Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val
35 40 45
Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys
50 55 60

<210> 40
<211> 4
<212> PRT
<213> human

<400> 40
Asn Phe Glu Val
1

<210> 41
<211> 5
<212> PRT
<213> human

<400> 41
Val Asn Phe Ala Val
1 5

<210> 42
<211> 5
<212> PRT
<213> human

<400> 42
Val Lys Val Asp Ala
1 5

<210> 43
<211> 5
<212> PRT
<213> human

<400> 43
Trp Lys Met Asp Ala
1 5

<210> 44
<211> 5
<212> PRT
<213> human

<400> 44
Val Lys Ala Asp Ala
1 5

<210> 45
<211> 5
<212> PRT
<213> human

<400> 45
Val Lys Lys Asp Ala
1 5

<210> 46
<211> 5
<212> PRT
<213> human

<400> 46
Val Lys Glu Asp Ala
1 5

<210> 47
<211> 5
<212> PRT
<213> human

<400> 47
Val Lys Ile Asp Ala
1 5

<210> 48
<211> 5
<212> PRT
<213> human

<400> 48
Val Lys Met Ile Ala
1 5

<210> 49
<211> 5
<212> PRT
<213> human

<400> 49
Val Lys Met Asn Ala
1 5

<210> 50
<211> 5
<212> PRT
<213> human

<400> 50
Val Lys Met Glu Ala
1 5

<210> 51
<211> 5
<212> PRT
<213> human

<400> 51
Val Lys Met Asp Glu
1 5

<210> 52
<211> 5
<212> PRT
<213> human

<400> 52
Val Lys Met Asp Lys
1 5

<210> 53
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 53
ggagaggata tcatggagcc agtagat 27

<210> 54
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 54
tacatggcgg ccgcctactt actgctttg 29

<210> 55
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 55
ggatgtgata tctttcttct tcagcatcac caagg 35

<210> 56
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 56
ctgagatatc aagctactgt cttctatcga acaagc 36

<210> 57
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 57
gcgcgatatc cccaccgtac tcgtcaattc c 31

<210> 58
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 58
ctgctgtggc ggccgcctag ttctgcatct gctc 34

<210> 59
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 59
tgccccgcgc ggccgcgcga tgctgcccgg 30

<210> 60
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 60
atggtgtggc ggccgcagac gccgctgtca cc 32

<210> 61
<211> 36
<212> DNA
<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 61
agcgcacaag cttccccgcg cagggtcgcg atgctg 36

<210> 62
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 62
ggatgtaagc tttttcttct tcagcatcac caagg 35

<210> 63
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 63
tgcagatatc ctggagccag tagatcctag ac 32

<210> 64
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 64
gctggatatc ctctgctttg atagagaagc 30

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 September 2003 (04.09.2003)

PCT

(10) International Publication Number
WO 2003/072041 A3

(51) International Patent Classification:
C07H 21/02 (2006.01) *C12N 1/20* (2006.01)
C12P 21/06 (2006.01) *C12N 15/00* (2006.01)

(21) International Application Number:
PCT/US2003/005458

(22) International Filing Date:
23 February 2003 (23.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/360,274 27 February 2002 (27.02.2002) US

(71) Applicant *(for all designated States except US)*: **MERCK & CO., INC.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ESPESETH, Amy, S.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **FERRER, Marc** [ES/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **FLORES, Osvaldo, A.** [CL/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **HAZUDA, Daria, J.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **INGLESE, James** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **MILLER, Michael, D.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **REGISTER, Bruce** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **SHI, Xiao-Ping** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **SIMON, Adam, J.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **ZUCK, Paul, D.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

(74) Common Representative: **MERCK & CO., INC.**; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
3 August 2006

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ASSAYS TO MONITOR AMYLOID PRECURSOR PROTEIN PROCESSING

(57) Abstract: The present invention provides DNA constructs, genetically engineered host cells, and methods for identifying inhibitors of amyloid precursor protein (APP) processing. The methods provide for the convenient identification, in a single assay, of inhibitors of β-secretase and γ-secretase as well as other forms of APP processing. The methods rely on fusion proteins of APP and transcription factors in which APP processing releases the transcription factors, allowing the transcription factors to activate transcription of a reporter gene. Inhibitors are identified as substances that block or diminish transcription factor release from the fusion protein, thereby causing a diminution of reporter gene readout.

WO 2003/072041 A3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/05458

A. CLASSIFICATION OF SUBJECT MATTER

IPC: C07H 21/02(2006.01);C12P 21/06(2006.01);C12N 1/20(2006.01),15/00(2006.01)

USPC: 536/23.1;435/69.1,252.3,320.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 536/23.1; 435/69.1, 252.3, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,811633 A (Haseltine et al) 1 September 1998 (01.09.1998), nucleic acid encoding HIV-1Tat transcription factor and trans-activating activity.	1-16
A	US 5,877,015 A (Hardy et al) 2 March 1999 (02.03.1999), SEQ ID NO:4 discloses APP695.	1-16
A	US 5,811,063 (Wadsworth et al) 22 September 1998 (22.09.1998), see claims disclosure of APP695.	1-16

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 March 2006 (20.03.2006)	01 JUN 2006
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner of Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	Authorized officer Janet Andres Telephone No. 308-1235

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/05458

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-16

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

PCT/US03/05458

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group 1 Claim(s) **1-16**, drawn to a DNA molecule and method of identifying a substance that inhibits APP processing.
Group 2 Claim(s) **17-20**, drawn to a method of identifying a substance that inhibits APP comprising a transcription factor fused to APP.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack unity of invention because they are not so linked as to form a single general inventive concept under PCT Rule 13.1. The species listed do not share a common structure or function.

The species are as follows:

The following claim(s) are generic:

Group 1, Claims 1, 4, 5, 6, 7, 13, 14, 15, 16, all APP species listed therein.

In order for more than one species to be examined, the appropriate additional examination fees must be paid.

Applicant is invited to select one species per generic claim listed per group elected. The inventions listed as Groups 1-2 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Group 1 is drawn to the special technical feature of a DNA molecule comprising APP695 fused to a transcription factor, which is not required by any of the other groups.

Group 2 is drawn to the special technical feature of a fusion protein comprising APP which is cleaved by both beta-secretase and gamma-secretase, which is not required by any of the other groups.